Cartel Formation under Incomplete Information

Pouyet, J.; Verouden, V.C.H.M.

Publication date:
2000

Citation for published version (APA):
Pouyet, J., & Verouden, V. C. H. M. (2000). *Cartel Formation under Incomplete Information: on the Requirements of Collusion-Proofness*. (CentER Discussion Paper; Vol. 2000-01). Microeconomics.

Center
for
Economic Research

No. 2000-01

CARTEL FORMATION UNDER INCOMPLETE INFORMATION: ON THE REQUIREMENT OF COLLUSION-PROOFNESS

By Jérôme Pouyet and Vincent Verouden

January 2000

ISSN 0924-7815

Cartel formation under incomplete information: On the requirement of collusion-proofness[*]

Jérôme Pouyet[†] and Vincent Verouden[‡]

Abstract

The existing literature on cartel formation uses the Revelation Principle to characterize the outcomes that a cartel can achieve in situations of incomplete information. In this paper we verify whether one assumption underlying its application, viz. that subgroups of firms do not attempt to jointly change the terms of the cartel in their favour, is not too restrictive. The principal result of the paper is that this is not the case in the context of industries where firms can be of two types (either efficient or inefficient): even when there are many firms in the industry, cartels can be formed that are not subject to internal manipulation.

JEL Classification: L41, D82.

Keywords: Cartel formation, asymmetric information, subcoalitions, collusion-proofness.

[*]We are very grateful to Jean-Jacques Laffont and Eric van Damme for helpful discussions. We also thank Patrick Rey for his comments and the participants at the working group on competition policy in GREMAQ.

[‡]GREMAQ (Toulouse) and CERAS, ENPC, 28 rue des Saints-Pères, 75007 Paris, France. E-mail: pouyet@enpc.fr.

[‡]CentER, Tilburg University, P.O. Box 90153, 5000 LE Tilburg, The Netherlands. E-mail: v.c.h.m.verouden@kub.nl.

1 Introduction

It is generally thought that the likelihood of firms forming a cartel is greater in concentrated industries than in industries with many firms. Not only because it is, so the argument goes, easier to monitor a cartel agreement in the relatively surveyable environment of a tight oligopoly (cartel enforcement argument) but also because it may be easier or more attractive for fewer firms to come to terms about the conditions applying to the cartel (cartel formation argument).

One element that can be a source of difficulty in the formation of cartels is the problem of *incomplete information* with respect to the cost levels of the participating firms. This information asymmetry may pose a problem at the stage where the cartel must determine the conditions of the cartel agreement (e.g. production quotas, fixed market shares) for the participants. Obviously, the conditions of the cartel agreement also bear on the decision to join the cartel in the first place. For example, a firm that is relatively efficient will typically only be satisfied with a production quota that somehow steps up to this fact (in other words, the quota must be relatively large), otherwise he may just prefer to compete on the market. This, however, should induce firms that are less efficient to overstate their efficiency in order to obtain a higher share of the cartel output. But when every firm is saying to be efficient (or saying to have become more efficient since the latest negotiations) and claiming large quota, this reduces the attractiveness of the cartel for firms that are effectively among the most efficient. An illustration of the ensuing difficulties is given by Eckbo (1976), who found that in a sample of international cartels that were temporarily successful but then broke down, almost half were ended due to internal squabbling over how to share the profits.

The extent to which cartel agreements can overcome the conflicting requirements mentioned above has been the subject of extensive research, cf. Roberts (1985), Kihlstrom and Vives (1989,1992) and Cramton and Palfrey (1990). In order to characterize the outcomes that a cartel can achieve in situations of incomplete information, these authors have approached the issue of cartel formation using a standard mechanism design approach: in a (Bayesian) Cournot industry, there is a 'cartel manager' who proposes a cartel arrangement and determines the optimal quotas depending on the costs each firm announces to have. Given this scheme, firms decide whether or not to join and, if they do, they announce their costs. A proposed cartel agreement is called 'efficient' when only the firm(s) with the lowest cost produce(s). In order to form such a cartel, the cartel manager must, according to the well-known Revelation Principle, propose a scheme (possibly involving side payments) that ensures individual participation (Bayesian individual rationality, BIR) and induces the firms to individually reveal their cost information (Bayesian incentive compatibility, BIC).

In a setting where the number of possible efficiency types is limited to two (firms are either efficient or inefficient), Kihlstrom and Vives (1989, 1992) have shown that the formation of an efficient cartel *is* possible, both in the case of a duopoly and in the borderline case of an industry comprising infinitely many firms (modelled as a continuum of firms). The reason of this latter, rather surprising, result is that in an atomic industry there is no uncertainty about the type of firm that should produce nor about the fraction of efficient firms being present, so that it turns out to be not so difficult to reconcile all individual participation and incentive requirements. This result holds for all meaningful probability distributions on the two cost types.

Cramton and Palfrey (1990), however, have shown that if the private information of the firms can take a continuum of values (ranging from being very efficient to being very inefficient), the first best outcome can only be implemented if there are not too many firms in the industry. With many firms, they show that it becomes increasingly difficult to reconcile all the individual incentive and participation requirements outlined above. In this sense, their main result is a confirmation of the general idea that it is easier to form a cartel in an industry with a few firms than in one with many. Nonetheless, the extent to which their result is driven by assuming the uniform probability distribution (assigning equal probability to all possible cost types, even the most extreme values) is not clear. Furthermore, assuming that the private information of the firms can take a continuum of values leads to the rather peculiar result that when less than unanimous consent is needed to ratify the cartel agreement (only a proportion $\alpha < 1$ of firms must agree with it), the impossibility result is lost. In other words, with less than unanimous consent, an efficient cartel can be implementable in industries with many participants[1].

The current literature, therefore, does seem to provide *some* justification for the generally held belief that cartels are most difficult to form in industries with many firms, but it fails to do so in several constellations, most notably in the context of a finite number of cost types with general probability distributions. In this paper we will again consider the issue of cartel formation in this latter context, but from a different angle: in order to characterize the possible outcomes that a cartel can achieve, we propose to explore the additional requirement of *collusion-proofness*. The above mentioned models of the cartel manager (the principal) trying to obtain the efficient cartel outcome by inducing truthful cost revelation by the firms (the agents) all use the standard assumption that every agent behaves non-cooperatively: no communication is possible between the agents, which is a standard assumption for the Revelation Principle. The aim of the current paper is to see whether the obtained results continue to hold when communication between groups of firms *cannot* be excluded and, in particular, when groups of firms try to (secretly)

[1] Cramton and Palfrey show that when less than unanimous consent is needed to ratify the cartel agreement, an efficient cartel is implementable, *especially* when the number of firms is very large. This is due to a special feature of the continuum type model, namely the fact that when the number of firms grows, the lowest cost firm that is supposed to produce the entire cartel output has an increasingly small measure (it is one of the many firms). As in the limit, it is only the lowest cost firm whose participation constraint cannot be met in the face of the incentive requirements with respect to the less efficient firms, this firm will not be able to stop the cartel from going ahead if a majority rule applies.

coordinate their cost announcements in order to obtain a better result.

So as to investigate the role of this possibility, we will model this kind of collusion following Laffont-Martimort (1997, 1998): in a first stage, the cartel manager (the principal) proposes a cartel contract ('grand mechanism') (q, t), where q is the vector of quantities to be produced by the cartel members and t the vector of side transfers as functions of the cost messages[2]. If any firm refuses this contract, the firms will compete à la Cournot. In phase 2, a 'mediator' (the third party) proposes to a group of firms a side contract (ϕ, y), where ϕ is the vector of manipulated cost messages of the considered group to be sent to the principal and y the vector of internal side payments. Finally, the firms in the group decide whether or not to accept the third party's program. If not, they will non-cooperatively send messages to the principal.

In the framework of Laffont and Martimort (1997, 1998), where, by assumption, collusion can only happen by all agents together, the implementable contracts can be characterized using the *Collusion-Proofness Principle.* This principle states that the cartel manager can, without loss of generality, restrict the cartel contract he proposes to be a mechanism such that no collusion takes place at the equilibrium. In our setting, where collusion may take place by coalitions of variable size, we will determine to what extent collusion-proofness is relevant.

The possibility of collusion by subcoalitions is shown to change the set of implementable rules, but *not* to change the principal result that efficient cartel formation is possible for any number of firms in the industry. Partly, this is due to the strong congruence of interest between the cartel manager and the individual cartel members: after all, the cartel manager is acting costlessly on behalf of the members, by maximizing their total expected profits. But there is more to it. Typically, the transfer schemes that are able to implement collusion-proof cartels appear to satisfy a property that we call *partial anonymity:* the cartel transfers to the inefficient firms are taken not to depend on the number of inefficient firms present in the cartel. The following intuition is then at the heart of the result: consider an inefficient firm that is instructed by the third party to represent itself to the cartel manager as being efficient. This firm will then get more revenue out of production, but will at the same time exert a negative externality on all firms that are truly efficient, including the efficient firms within the subcoalition (as the cartel output must now be divided over a larger number of firms). If, furthermore, the cartel transfers are partially anonymous, then the effect of the contemplated action is neutral with respect to to the other inefficient firms in the subcoalition. So we are left to compare the direct effect on the pay-off of the particular inefficient firm with the negative externality on the efficient firms in the subcoalition. It follows that as soon as the inefficient firm is made individually indifferent between lying and truthtelling, that then the group incentive constraints

[2]These transfers will, for simplicity, be considered as monetary transfers. In principle, however, they can stand for any kind of compensation, e.g. promising not to enter a different geographical market where the other firms are already active.

with respect to downward cost announcements will be satisfied. This line of reasoning identifies a transfer scheme that enables the implementation of cartels that are not subject to internal manipulation, in industries of any size. Finally, we identify cases in which the extra requirement of collusion-proofness need not have an impact on the minimal level of transfers that is required to form these cartels.

2 The model

We consider the possibilities for cartel formation in an industry with n firms producing a homogeneous good. Let q^i denote the production level of firm $i \in N = \{1, ..., n\}$ and $Q = \sum_{i=1}^{n} q^i$ the total market production. As in Kihlstrom and Vives (1989, 1992) and Cramton and Palfrey (1990), market demand for the good is given by a linear inverse demand function $p(Q) = 1 - Q$. The firms in the industry can be of two types: either efficient (their unit cost of production is low, $\underline{c}$) or inefficient ($\overline{c}$). Let Δc denote the cost difference $\overline{c} - \underline{c}$. The levels of efficiency are private knowledge to the firms, they are not observable to other firms or to the cartel (embodied by the cartel manager). Nonetheless, it is common knowledgde that the probability of an individual firm being efficient is $\mu \in (0, 1)$ and that the types are independently and identically distributed (i.i.d.). The output levels of the individual firms are observable.

The cartel manager proposes the firms a cartel contract $\{q_i, t_i\}_{i \in N}$, where $q_i = q_i(\tilde{c}^N)$ is the vector of quantities to be produced by the cartel members and $t_i = t_i(\tilde{c}^N)$ the vector of transfers as functions of the cost reports $\tilde{c}^N$. The transfers are restricted to be budget balanced in each state of nature.

The gain of firm i is given by

$$\pi^i(\tilde{c}^N) = [p(Q(\tilde{c}^N)) - c^i]q^i(\tilde{c}^N) - t^i(\tilde{c}^N).$$

The objective of the cartel manager is to form a cartel that is *efficient*, as defined as follows:

Definition 1 *An efficient cartel is a cartel that maximizes the industry profit* $\sum_{i \in N} \pi^i(\tilde{c}^N)$ *in each state of nature.*

A proposed cartel agreement is efficient when only the firm(s) with the lowest cost produce(s) and the total industry output equals the monopoly output for this cost level, being $q^m(c) = \arg\max_q \{(p(q) - c)q\}$. A symmetric and efficient cartel amounts to the following rule chosen by the cartel manager

$$q^i(\tilde{c}^N) = \begin{cases} \frac{1}{j} q^m(\underline{c}) & \text{if } \tilde{c}^i = \underline{c}, \\ 0 & \text{if } \tilde{c}^i = \overline{c} \text{ and } j \neq 0 \\ \frac{1}{n} q^m(\overline{c}) & \text{if } \tilde{c}^i = \overline{c} \text{ and } j = 0 \end{cases}$$

where j is the total number of firms that report to be efficient to the cartel manager[3]. Let $t(j)$ be the total transfer to be paid by the efficient firms (and consequently to be received by the inefficient firms). Then, we have

$$t^i(\tilde{c}^N) = \begin{cases} -\frac{1}{j}t(j) & \text{if } \tilde{c}^i = \underline{c}, \\ \frac{1}{n-j}t(j) & \text{if } \tilde{c}^i = \overline{c}. \end{cases}$$

Notice that when every firm reports to be efficient ($j = n$) or every firm reports to be inefficient ($j = 0$), no transfers are occurring: $t(n) = t(0) = 0$. For the rest of the paper, we will use the following conventions: all the transfers $t(j)$ are non-negative. This means that, unless the transfers are zero, the efficient firms will pay monetary transfers to the cartel manager, which are then redistributed to the inefficient firms. This is an innocuous assumption because we are primarily interested in efficient cartels in which only the most efficient firms produce and consequently derive gains from production. In order that the inefficient firms are willing to take part in the cartel, they must normally be compensated for not producing.

3 The individual constraints

When the firms cannot communicate among each other (i.e. they adopt a Bayesian-Nash behaviour with respect to the actions of the other firms), then the Revelation Principle applies[4]. In order to characterize the set of attainable outcomes of the cartel, we can restrict ourselves to direct mechanisms that satisfy a set of Bayesian incentive compatibility constraints (called henceforth BIC) in order to obtain the truthful revelation of the private information held by each firm. But the cartel manager must also ensure that the firms are willing to participate in the cartel: this gives another set of constraints, the Bayesian individual rationality constraints (called BIR).

3.1 The Bayesian incentive constraints

Let us call $\mathbf{E}_j\{t(j)\} = \sum_{j=0}^{n} \binom{n}{j}\mu^j(1-\mu)^{n-j}t(j)$ the expected transfer to be given by the efficient firms to the inefficient firms. For a symmetric cartel, we can represent the Bayesian incentive constraints of the two types of firms simply by one lower bound and one upper bound on this expected transfer. Obviously, the lower bound refers to the Bayesian incentive compatibility constraint of an inefficient firm: it gives the minimum expected transfer that the inefficient firms must receive to truthfully reveal their information to the cartel manager. In the case of

[3] In a symmetric equilibrium, only the private information is relevant: to characterize the state of nature it is enough to know the number of efficient (or inefficient) firms in the cartel.

[4] For the Revelation Principle, see Green and Laffont (1979) and Baron and Myerson (1982), among others.

an efficient cartel, this constraint is

$$BIC(\overline{c}) : \mathbf{E}_j\{t(j)\} \geq n\mu(1-\mu)[\frac{1-(1-\mu)^n}{\mu n}(\pi^m(\underline{c}) - \Delta c q^m(\underline{c})) - \frac{1}{n}(1-\mu)^{n-1}\pi^m(\overline{c})].$$

Note that when $\pi^m(\underline{c}) - \Delta c q^m(\underline{c}) \leq 0$, an inefficient firm has no interest to announce it is efficient, because it derives a negative gain from producing the low cost monopoly quantity and it must also pay some transfers to the firms that do announce to be inefficient. Hence, in this case there is no problem of information revelation from the point of view of an inefficient firm. Therefore, we will restrict ourselves to the case in which $\pi^m(\underline{c}) - \Delta c q^m(\underline{c}) \geq 0$.

In a similar fashion, the incentive constraint of an efficient firm gives a upper limit to the expected transfer to be given by the efficient firms:

$$BIC(\underline{c}) : \mathbf{E}_j\{t(j)\} \leq n\mu(1-\mu)[\frac{1-(1-\mu)^n}{\mu n}\pi^m(\underline{c}) - \frac{1}{n}(1-\mu)^{n-1}(\pi^m(\overline{c}) + \Delta c q^m(\overline{c}))].$$

3.2 The participation constraints

Following Kihlstrom and Vives (1989, 1992), we suppose that if at least one firm refuses to participate in the cartel then the cartel breaks down and the firms play a standard Cournot competition game under asymmetric information. At this stage, we must be explicit about the beliefs of each firm following the rejection of the cartel contract. We assume that if firm i does not accept the cartel contract then this firm does not change its beliefs on the types of the $n-1$ remaining firms, nor do the other firms change their beliefs on the type of the firm that has refused the cartel contract. This is the assumption of passive beliefs as made by Kihlstrom and Vives[5].

In a symmetric and efficient cartel, we can also express the Bayesian individual rationality

[5]Whether the assumption of passive beliefs is crucial remains an open question in this model with two types. If efficient firms accept the symmetric cartel contract proposed by the cartel manager but the cartel breaks d own, it must be the case that an inefficient firm has rejected this contract. The efficient firms know now that there is at least one inefficient firm in the market. This should increase the competitive pressure for the inefficient firms (as the efficient firms will be producing more relative to the prior situation) and therefore relax their participation constraints. The reverse holds for an efficient firm: its participation constraint is in fact more tight than the one implied by passive beliefs. We plan to explore this issue in further research. We also refer the reader to Cramton and Palfrey (1995) for an analysis of the effects of learning from disagreement with the concept of ratifiability. In any case, the assumption is necessary in order to have a tractable analysis of collusion at the level of the subcoalitions.

constraints of an efficient and an inefficient firm as bounds on the expected transfers[6],

$$BIR(\underline{c}) : \mathbf{E}_j\{t(j)\} \leq n\mu[\frac{1-(1-\mu)^n}{\mu n}\pi^m(\underline{c}) - \mathbf{E}\{\pi_{\text{Cournot}}(\underline{c})\}$$

and

$$BIR(\overline{c}) : \mathbf{E}_j\{t(j)\} \geq n(1-\mu)[\mathbf{E}\{\pi_{\text{Cournot}}(\overline{c})\} - \frac{1}{n}(1-\mu)^{n-1}\pi^m(\overline{c})].$$

3.3 Individual implementability

From now on we will use the following terminology.

Definition 2 *An efficient cartel is individually implementable if there exists a set of transfers such that all the individual constraints (participation and incentive compatibility) are satisfied*[7].

An illustration of the individual incentive compatibility constraints and the participation constraints, as defined previously, is given in figure 1 for the case of three firms.

The fact that in the present context (of a finite number of types, the probability distribution of which is i.i.d.), budget-balanced transfers can be found that satisfy the Bayesian incentive compatibility constraints for both types is in fact just an illustration of the result of d'Aspremont and Gérard-Varet (1979). The fact that is possible to find transfers that satisfy the participation constraints for both types (the upperbound of $BIR(\underline{c})$ is above the lowerbound of $BIR(\overline{c})$) is not very surprising either: as the cartel profit exceeds the competitive industry profit in each state of nature, it will be possible to form the cartel, given that every firm tells the truth.

Comparing the lowerbound of $BIC(\bar{c})$ with the upperbound of $BIR(\underline{c})$ and the lowerbound of $BIR(\bar{c})$ with the upperbound of $BIC(\underline{c})$ is less straightforward. In particular the first comparison poses analytical problems. Nonetheless, we obtain the following proposition.

Proposition 1 *Sufficient conditions for the cartel to be individually implementable are that (i)* $\Delta c \leq \Delta c_{ES}$*, where* $\Delta c_{ES} = \Delta c_{ES}(\mu, n)$ *is the largest cost difference for which efficient firms are still willing to share the cartel profits equally with the inefficient firms and (ii)* $\Delta c \geq \Delta c^0_{BIC}$*,*

[6] For the Cournot equilibrium we must distinguish between two cases: one in which both efficient and inefficient firms are producing and one in which only efficient firms produce. Indeed, if the probability of being efficient (μ), the number of firms in the market (n), or the cost differential (Δc) are large, then an inefficient firm will anticipate a high competitive pressure and will prefer not to produce at the Cournot equilibrium under incomplete information. In that case, only efficient firms will be active in the market and make profits equal to $\frac{(1-\underline{c})^2}{(2+(n-1)\mu)^2}$. If the inefficient firms also produce at the Cournot equilibrium, the profit of a c-type firm is given by $\frac{1}{4}[1-c-\frac{n-1}{n+1}(1-\mathbf{E}\{c\})]^2$ where $\mathbf{E}\{c\} = \mu\underline{c} + (1-\mu)\overline{c}$ is the expected cost.

[7] Our definition concerning the implementability of an efficient cartel includes the incentive constraints and the individual participation constraints. Usually, the term 'implementability' refers only to the problem of information revelation. But we use this terminology to make a clear-cut distinction between the individual constraints and the constraints concerning the subcoalitions that we will define later on.

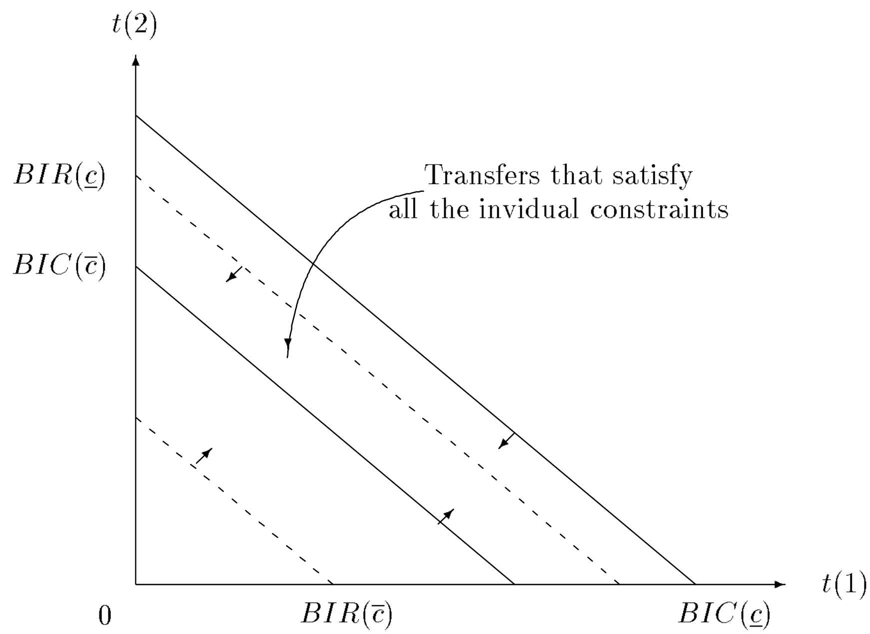

Figure 1: Individual implementation of an efficient cartel is always possible.

where $\Delta c^0_{BIC} = \Delta c^0_{BIC}(\mu, n)$ *is the level of* Δc *for which, even in the absence of transfers, the inefficient firms are indifferent between lying and truthtelling.*

Under the first condition, low cost firms are effectively willing to share the profits from the cartel equally with the high cost firms. Under this transfer scheme, the 'equal sharing rule', high cost firms have no incentive to misrepresent themselves, so that $BIC(\overline{c})$ is satisfied. This sufficient condition obviously amounts to cost differences that are small. The second condition stems from the fact that $BIC(\overline{c})$ can already be satisfied by *zero* transfers, when the cost differences are large. In that case the bound imposed by $BIC(\overline{c})$ (being zero) is clearly below the bound of $BIR(\underline{c})$ (positive).

As in Cramton and Palfrey (1990), it turns out to be impossible to completely compare the two constraints $BIR(\underline{c})$ and $BIC(\overline{c})$ for general parameters without using simulations. For cost differences inbetween 'small' and 'large', we resort to simple simulations that show that the cartel is also implementable for these parameter values[8]. The result is in line with the findings of Kihlstrom and Vives (1989, 1992) for the polar cases of two firms and a continuum of firms.

[8]The simulations are done after first having reduced the relevant inequality (which is an inequality in four parameters: $\underline{c}$, $\overline{c}$, μ and n), into an inequality with only two parameters (μ and n); see the appendix for the details.

Note that there exists in fact a whole region (of non-zero measure) of transfers that enable the cartel manager to simultaneously satisfy the Bayesian incentive constraints and the Bayesian participation constraints for both types of firms. These constraints can be binding or not, depending on the profile of transfers chosen.

3.4 Individual implementability with minimal transfers

Having established the result that an efficient cartel can be implemented, it is interesting to focus on a few characteristics of the possible cartels. One such characteristic is the minimal transfers that are necessary to operate it. After all, a cartel that operates in an industry that is surveyed by an antitrust authority will normally, in order to minimize the risk of detection, have a certain preference for cartel schemes that involve as few transfers as possible. We find, on the basis of a sufficient condition and simulations, that whenever positive expected transfers are necessary to implement the efficient cartel, it is the Bayesian incentive compatibility constraint of an inefficient firm that determines the level of the necessary expected transfers.

In terms of figure 1, this result tells us that whenever positive transfers are needed, the lowerbound of $BIC(\bar{c})$ is above the lowerbound of $BIR(\bar{c})$. Note that this is fairly intuitive: when the cost difference is almost zero ($\bar{c} \approx \underline{c}$), high transfers will be necessary to refrain a 'high cost' firm from telling that it is low cost; in fact, a high cost firm will require nothing less than equal sharing of the cartel profits among all firms. On the other hand, $BIR(\bar{c})$ implies that a high cost firm gets more than what it would get under Cournot competition under (almost) complete information and (almost) identical costs. Obviously, the equal sharing pay-off exceeds this profit: hence, for $\bar{c} \approx \underline{c}$, the lowerbound of $BIC(\bar{c})$ is above that of $BIR(\bar{c})$. When the cost difference increases, the cost of lying increases and the lowerbound of $BIC(\bar{c})$ goes down – but so will the lowerbound of $BIR(\bar{c})$, as the outside option of Cournot competition becomes less attractive.

It follows that, when positive, the minimal required expected transfer is equal to

$$n\mu(1-\mu)[\frac{1-(1-\mu)^n}{\mu n}(\pi^m(\underline{c}) - \Delta c q^m(\underline{c})) - \frac{1}{n}(1-\mu)^{n-1}\pi^m(\bar{c})]$$

i.e. the right hand side of the Bayesian incentive compatibility constraint for an inefficient firm given above. The influence of the industry parameters on the minimal required transfer can be inferred from this value: for given n, the larger the cost difference (Δc) and the smaller the probability of an other firm being efficient (μ), the smaller the expected transfers can be. The intuition of this result is straightforward. Given that an inefficient firm only gets to produce when a $\bar{c}$-cartel is formed, the gain of truthtelling increases with the probability that this cartel actually forms. This probability is equal to the probability that there is *no* efficient firm around, $(1-\mu)^n$. Hence, the smaller the probability that the other firms are efficient, the better for a firm that announces to be inefficient. As regards the cost difference Δc, the larger this difference, the less attractive the option of lying becomes to the inefficient firm. An inefficient firm, if it announces to be efficient, will be in a position to produce but this will be all the more costly, the higher its unit cost of production relative to the cost level of the efficient firms in the cartel.

Is it possible for a cartel to operate without the use of any transfers at all? For this to be possible it must be that the transfer scheme $t(j) = 0, \forall j$ (the 'zero transfer rule') satisfies the incentive and participation constraints for both efficient and inefficient firms. Obviously, the efficient firms will be very contented with the zero transfer rule: they get to produce but do not have to give any of their profits to the firms announcing to be inefficient. Both the incentive constraint and the participation constraint of the efficient type are satisfied. With respect to the inefficient firms, we know from the previous discussion that we only need to see whether the zero transfer rule satisfies the corresponding incentive compatibility constraint: if incentive compatibility is ascertained for an inefficient firm under the zero transfer rule then it must be the case that the cost difference is quite large. But then the expected profits at the Cournot-Nash equilibrium are that small that participation by the inefficient firm is ensured as well. It follows that transfers are not necessary to operate the cartel when the cost difference is sufficiently large or when the probability of an other firm being efficient, is sufficiently small.

4 The formation of subcoalitions

As we have seen in the previous section, there is a lot of leeway for the cartel manager to implement the monopoly outcome. Indeed, there is a whole region of transfers that can be used for this purpose (or a hyperplane, in case the cartel manager prefers to use minimal transfers). However, depending on the chosen transfers, it may be worthwile for some subcoalitions of firms to come together and try to manipulate the cartel outcome, in order to reach higher gains from participation in the cartel.

4.1 The stake of collusion: a heuristic presentation

Assume, as a preliminary step to the forthcoming analysis, that a subcoalition of two inefficient firms can overcome their informational problem and can form (for example, this is the case if the firms in the subcoalition can credibly disclose their private information to each other or if they have a technology to communicate with each other). Suppose that this subcoalition is also able to use side transfers between its members and that it tries to maximize the gains of its members. Considering again the example of three firms, the expected total gain of the subcoalition composed of, say, two inefficient firms when they reveal truthfully their costs to the cartel manager is $\mu t(1) + (1-\mu)\frac{1}{3}\pi^m(\overline{c})$. If these two inefficient firms manipulate their

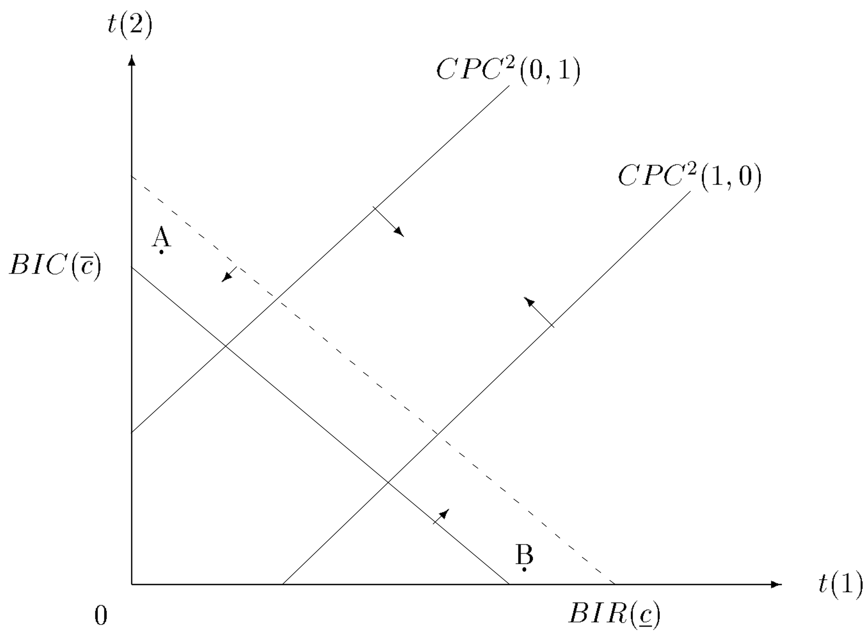

Figure 2: The stake of collusion with three firms.

announcements and claim to the cartel manager that they are composed of one efficient and one inefficient firm then the expected gain of the subcoalition becomes $\mu[\frac{1}{2}(\pi^m(\underline{c}) - \Delta c q^m(\underline{c})) + \frac{1}{2}t(2)] + (1-\mu)[\pi^m(\underline{c}) - \Delta c q^m(\underline{c}) - \frac{1}{2}t(1)]$. In the case where the cartel manager chooses a set of transfers with a high $t(2)$ and a low $t(1)$, the subcoalition of two inefficient firms will have interest to coordinate their cost messages and to lie to the cartel manager. This leads to some inefficient production of the cartel output, which is not desirable from the point of view of the cartel manager. Requiring that the subcoalition of two (2) firms reveals truthfully that it consists of zero (0) efficient firms instead of announcing to contain one (1) efficient firm gives a collusion-proofness constraint, $CPC^2(0,1)$. For example, in Figure 2, points A and B enable the cartel manager to implement individually an efficient cartel but do not resist to the formation of active subcoalitions. Note that there are four other constraints preventing deviations of the remaining subcoalitions. This highlights an interesting property of the model we study. Contrary to the previous models on collusion under asymmetric information, there can be many different stakes of collusion depending on the transfers that are used by the cartel manager to individually implement an efficient cartel.

In the next subsection, we model how collusion can take place. We depart from the crude modelling device used in the previous example, where firms can freely share their soft information, and explicitly take into account the informational problem at the coalition level.

4.2 Subcoalition formation under incomplete information

In this subsection, we propose to use the framework of Laffont and Martimort (1997, 1998) to model the formation of subcoalitions within the cartel. We first begin with the timing that explains how and when a third party can propose to a given subcoalition of k firms a side mechanism aiming at manipulating the reports to be sent by the firms in the cartel contract proposed by the cartel manager.

From now on, we will use the following notation. We denote S for the set of indices of the k firms that belong to the subcoalition and $N\backslash S$ for the set of indices of the $n-k$ other firms that belong to the cartel but not to the subcoalition. Accordingly, c^S (resp. $c^{N\backslash S}$) is the vector of the true cost parameters of the firms belonging to the subcoalition (resp. of the firms belonging to the cartel but not to the subcoalition). The timing is as follows:

1. Nature draws $c^i \in \{\underline{c}, \overline{c}\}$, the private information of firm $i \in N$ according to the common knowledge probability distribution $\{\mu, 1-\mu\}$. Each firm only learns its own type.

2. The cartel manager proposes all the firms a cartel contract, as defined before.

3. Firm i decides to accept or reject the cartel contract, for $i \in N$. If at least one firm decides to reject it, then the cartel breaks down and the firms come back to Cournot competition under asymmetric information.

4. If all the firms accept the cartel contract then a third party proposes a side mechanism to the firms $i \in S$. This mechanism is composed of a manipulation function $\phi : \{\underline{c}, \overline{c}\}^k \longrightarrow \{\underline{c}, \overline{c}\}^k$ of the messages to be sent to the cartel manager by the firms of the subcoalition, and a vector of monetary side transfers $y : \{\underline{c}, \overline{c}\}^k \longrightarrow \Re$ which must be budget-balanced in each state of nature. If one firm refuses the side mechanism, then the cartel contract is played non-cooperatively by the firms of the subcoalition. If all the firms of the subcoalition accept the side mechanism then they report non-cooperatively their private information to the third party. The third party then assigns the messages to be sent to the cartel manager by the firms in the subcoalition and promises to enforce the corresponding side-transfers.

5. Finally, the reports are sent by the firms into the cartel contract, the quantities and transfers proposed by the cartel manager in the cartel contract take place, as well as the side transfers proposed by the third party in the side mechanism[9].

[9]In the game the non-colluding firms play the standard game (they do not know that some other firms are colluding). Colluding firms know that they are colluding. The equilibrium that we will derive is therefore the Bayesian equilibrium of an inconsistent game, not of a Harsanyi's game.

The objective of the third party is to maximize the sum of the expected gains of the colluding firms. What can the third party achieve, given that he is uninformed about the private information of the firms in the subcoalition? Because the Revelation Principle applies at this stage of the game, there is no loss of generality in restricting the set of side mechanisms the third party proposes to the set of direct side mechanisms in which each firm of the subcoalition reveals truthfully its own private information[10]. Hence, the third party must satisfy a set of Bayesian incentive compatibility constraints. But the third party must also ensure that the firms of the subcoalition are willing to participate: a set of Bayesian individual rationality constraints must be verified by the third party as well, where the reservation value of one firm is its gain when the cartel contract is played non-cooperatively.

For a given state of nature c^S, we denote by $\underline{S}$ and $\overline{S}$ the set of efficient and inefficient firms. We find the following optimality conditions for the manipulation function of the third party:

$$\begin{aligned}\phi(c^S)^{opt} \in \arg\max_{\phi(c^S)} \mathbf{E}_{c^{N\setminus S}}\{\sum_{i\in\underline{S}} \pi^i(\underline{c}, \phi(c^S), c^{N\setminus S}) + \underline{\epsilon}_i \frac{1-\mu}{\mu}\Delta c q^i(\phi(c^S), c^{N\setminus S}) \\ + \sum_{j\in\overline{S}} \pi^j(\overline{c}, \phi(c^S), c^{N\setminus S}) - \overline{\epsilon}_j \frac{\mu}{1-\mu}\Delta c q^j(\phi(c^S), c^{N\setminus S})\}.\end{aligned}$$

The two variables $\underline{\epsilon}_i$ and $\overline{\epsilon}_j$ reflect the fact that there is an asymmetry of information between the members of the subcoalition. The cost of bridging this informational gap, embodied in $\underline{\epsilon}_i$ and $\overline{\epsilon}_j$, can prevent the subcoalition from realizing all the potential gains of the collusion. The values of the ϵs depend on whether the constraints in the program of the third party are binding or not[11].

5 Collusion-Proofness

Being faced with the possibility of collusion by a group of firms, what is the optimal response for the cartel manager? In related, but different, contexts, Laffont and Martimort (1997, 1998) have shown that the so-called *Collussion-Proofness Principle* applies. This principle states that there is no loss of generality to restrict the principal to use a grand mechanism that is collusion-proof.

[10] We assume that each third party takes in its collusive program only the individual incentive constraints into account. For the case of $n = 3$ firms this is the correct as all the subcoalitions are composed of just $k = 2$ individual firms. The third party for the two firms has to ensure that each firm reveals its piece of information. However, for the case of $n > 3$ firms we would in fact also have to consider the possibility that a sub-subcoalition could try to cheat on the third party. Clearly this would add more constraints in the program of the third-party, so that his intermediation would at best be equally efficient as the outcome implied by the constraints in our model. Our assumption therefore still helps us to determine an upper bound of the gain of the subcoalition

[11] When a given cartel contract amounts to ϵs being zero, the third party does not face problems having the information revealed and can implement the efficient side contract. However, it would be wrong to identify a situation in which $\epsilon = 0$ with a situation of complete information at the level of the third party. If this were the case, the cartel manager could incorporate this fact in its cartel contract and take advantage of it.

Definition 3 *A cartel contract proposed by the cartel manager is collusion-proof with respect to a subcoalition S if and only if the null side-mechanism $\{\phi = Id_{\{\underline{c},\overline{c}\}^k}, \{y^i = 0\}_{i\in S}\}$ is the optimal response for the third party.*

Differently stated, a grand mechanism is collusion-proof when the third party finds it optimal not to distort the announcements sent by the firms in the subcoalition to the cartel manager and when no monetary side-transfers take place between these firms.

Let us assess whether the Collusion-Proofness Principle applies to the current context of cartel formation as well. The common element in the settings of Laffont and Martimort (1997, 1998) is that the principal is, by assumption, only faced with one possible group of agents that is considering to collude, namely the entire group of agents. In this case, the logic that underlies the Collusion-Proofness Principle is very similar to the one of the Revelation Principle. The outcome of any general mechanism that triggers collusion can be replicated by a direct mechanism that neutralizes the incentive to misrepresent from the coalition's point of view. This allows for a clear characterization of the implementable allocations.

With the possibility of subcoalitions, it is straightforward to see that the Principle also applies if the cartel manager knows the size and identity of the coalitions possibly colluding[12] (we assume that one firm cannot be part of two subcoalitions at the same time). By contrast, when the cartel manager does not know which coalition(s) actually form – the setting of the current paper – things are different. For, if the cartel manager proposes a general cartel contract and one group colludes, a particular allocation will result, whereas if another group colludes, another allocation will result. Furthermore, a mechanism that is collusion-proof with respect to one subcoalition does not need to be collusion-proof with respect to another subcoalition. This implies that the Collusion-Proofness Principle cannot be generalised in the sense that the cartel manager can, without loss of generality, restrict attention to direct mechanisms that are collusion-proof with respect to to all conceivably active coalitions: it is not guaranteed that such mechanisms replicate the allocation of the general mechanisms[13]. In particular, it may be optimal for the cartel manager to focus on some groups and to tolerate possible collusion by other groups.

[12]Knowing the size and identity of the coalitions possibly colluding is not the same as knowing that they actively collude. The first part only refers to the coalitions being known to possess a technology to collude (i.e. a third party). As to the second part: even if the firms of the subcoalition know whether they have accepted the side contract or not, any attempt by the cartel manager to elicit this private information can be annihilated by the third party which can punish the deviators from the collusive side-mechanism as much as the cartel manager can reward those who reveal that they are colluding. This also implies that the cartel manager has to satisfy the individual constraints.

[13]This applies not only when the size k is unknown but also when k were known, but not the exact identity of the participating firms. Note furthermore that again this information cannot be elicited, not from the firms actually proposed a collusive side contract nor from the firms not being proposed such a contract.

Nonetheless, the focus on efficient cartels allows for the formulation of necessary and sufficient conditions for implementability. For, an efficient cartel can only be implemented if it can be ensured that no subcoalition finds it in its interest to overstate its efficiency (announcing that it contains more efficient firms than is the case in reality) as this would lead to inefficient production[14].

Let us call $CPC^k(l, l')$ the collusion-proofness constraint that prevents a subcoalition of size k composed of l efficient firms to lie and to annouce to the cartel manager that it is composed of l' efficient firms.

In the next section we will exhibit cartel contracts that implement the efficient cartel. In fact, these cartel contracts turn out to prevent *all* conceivable types of collusion by coalitions (that is, also collusion in the form of announcing to consist of more inefficient firms than is really the case). Whether there exist cartel contracts that satisfy the necessary and sufficient conditions for the implementation of an efficient cartel, but that do not prevent all types of collusion, remains an open question.

6 Implementability when collusion is possible

In this section, we ask the following question: is it possible to find a set of transfers such that an efficient cartel is implementable even in the presence of colluding coalitions? For expositional purposes, we will first address the question when it is not necessary at all to use transfers.

6.1 The zero transfer rule

As has already been said, the zero transfer rule is particularly appealing because it enables a cartel to form without having to proceed to any compensatory transfers. What is more, is that appears that whenever a symmetric and efficient cartel is individually implementable with the zero transfer rule, it also satisfies the relevant collusion-proof constraints. ($CPC^k(l, l')$, with $l < l'$).

The intuition is as follows. Let us start with a subcoalition of size k, consisting of l efficient firms and $(k - l)$ inefficient firms. Consider, first, the consequences (in terms of pay-offs for the subcoalition) of one inefficient firm being instructed by the third party to tell the cartel manager that it is efficient. These consequences can be divided into three parts: the change in revenue for the particular inefficient firm itself, the change for the l efficient firms in the subcoalition and the change for the $(k - l)$ inefficient firms. Now, under the zero transfer rule, inefficient firms never get transfers. Hence, for the inefficient firms it does not matter whether one of them is

[14]Note that we are considering symmetric cartel contracts. This is without loss of generality, given that the cartel manager does not know which coalition is formed.

instructed to lie. For the efficient firms, the consequences are undoubtly negative: as soon as the inefficient firm lies, there will be more 'efficient' firms in the cartel, so that the cartel output (i.e. the monopoly output) must be divided among more firms. In that case their pay-offs will be less than when the inefficient firm tells the truth. Hence, in lying the inefficient firm exerts a negative externality on the efficient firms in the subcoalition. So, we must compare the direct gain of lying for the inefficient firm with the negative externality for the efficient firms in the subcoalition. Let us clarify this comparison by means of the following extreme cases. For each case, we will write down the constraints to be satisfied by the zero transfer rule and see which constraint is the most demanding.

The case of $\mu \to 0$ In this case, the subcoalition will act under the assumption that there will be (almost) no efficient firms among the firms that are not in the subcoalition. Then an efficient cartel can be formed without transfers if

$$\begin{array}{rl} CPC^k(l,l') & \dfrac{l'-l}{l}\Delta c q^m(\underline{c}) \geq 0 \\ CPC^k(0,l') & \dfrac{k}{n}\pi^m(\overline{c}) - (\pi^m(\underline{c}) - \Delta c q^m(\underline{c})) \geq 0 \\ BIC(\overline{c}) & \dfrac{1}{n}\pi^m(\overline{c}) - (\pi^m(\underline{c}) - \Delta c q^m(\underline{c})) \geq 0. \end{array}$$

One can observe that whenever $BIC(\overline{c})$ is satisfied, $CPC^k(0,l')$ is also satisfied; the latter is more easily satisfied because a coalition takes into account the above mentioned externality, whereas a firm acting on its own does not. Furthermore, $CPC^k(l,l')$ is always satisfied: as there are no outside efficient firms, the coalition with $l \geq 1$ efficient firms prefers to have the whole monopoly output produced by its efficient firms only.

The case of $\mu \to 1$ Now, the subcoalition will anticipate that all firms are efficient; it is hence with certainty that the low cost cartel will form. An inefficient firm or a coalition with only inefficient firms will then have a strong incentive to misrepresent. The following constraints are to be met

$$\begin{array}{rl} CPC^k(l,l') & \dfrac{l'-l}{(l'+n-k)(l+n-k)}(-(n-k)\pi^m(\underline{c}) + (l+n-k)\Delta c q^m(\underline{c})) \geq 0 \\ CPC^k(0,l') & -\dfrac{l'}{l'+n-k}(\pi^m(\underline{c}) - \Delta c q^m(\underline{c})) \geq 0 \\ BIC(\overline{c}) & -\dfrac{1}{n}(\pi^m(\underline{c}) - \Delta c q^m(\underline{c})) \geq 0. \end{array}$$

The intuition that a coalition with only inefficient firms has a strong incentive to misrepresent, is reflected by the fact that the gain from lying with respect to $CPC^k(0,l')$ is larger than the gain of lying with respect to $CPC^k(l,l')$. Further, one can note that among the $CPC^k(l,l')$, it depends which of the constraints is most demanding. When k is small, the gain of lying is largest for $l' = k$; when k is large, the cost of lying is smallest for $l' = l + 1$.

The above two extreme cases, $\mu \to 0$ (all outside firms are inefficient) and $\mu \to 1$ (all outside firms are efficient) gave the result that whenever the zero transfer rule satisfies $BIC(\bar{c})$, it also satisfies all collusion-proofness constraints. This leads us to think that also for intermediate cases (in expectation, some outside firms are efficient, some are not) the same applies. Given the complicated character of the expressions for the general case, we are not able to prove this result in general. Nonetheless, we have

Proposition 2 *Suppose that only subcoalitions comprising a limited number of firms ($k \leq 2\sqrt{n}$) can form. Under the zero transfer rule, when an efficient cartel is individually implementable, it also implementable when collusion is possible.*

The reason is that only when k is not too large, we are able to rank the constraints $CPC^k(0, l')$ and $CPC^k(l, l')$ and to say that the former implies the latter. This is illustrated by the above exposition of the extreme cases. For most cases, it holds that whenever $CPC^k(0, l')$ is satisfied, $CPC^k(l, l')$ is satisfied as well. But when $\mu \to 0$ and Δc takes on the largest relevant value, the cost of lying is greater with respect to $CPC^k(0, l')$ than for $CPC^k(l, l')$ when k is large[15]. Simulations[16]confirm the intuition that the proposition extends to the case of large k as well.

Note that, under the zero transfer rule, also all $CPC^k(l, l')$ with $l' < l$ are satisfied: an efficient firm has no interest to lie because it always gets a part of the monopoly profit and does not have to give back anything to the inefficient firms[17].

6.2 Positive transfers

The intuition obtained from the zero transfer rule carries over to the case where individual implementabilty requires positive transfers. Let us again consider a subcoalition of size k, consisting of l efficient firms and $(k - l)$ inefficient firms, contemplating to announce that it consists of $l + 1$ efficient firms. The consequences of this action can as before be divided into three parts: the change in revenue for the particular inefficient firm itself, the change for the l

[15]This is so, because at this cost level, an inefficient firm that is instructed to lie makes no profit by producing the efficient quantity. Then, for a coalition without any efficient firms, lying simply means earning nothing and giving up its part of the inefficient cartel profit ($\frac{k}{n}\pi^m(\bar{c})$), an amount that is increasing in k. For a coalition with at least one efficient firm among its members, lying means reducing the quantity produced by truly efficient firms and giving up the concommittant profits.

[16]Available upon request.

[17]Given that all collusion-proofness constraints are satisfied, $\phi = Id$. Then, the incentive compatibility constraints of the third party program coincide with the incentive compatibility constraints of the cartel manager's program. Hence, if $BIC(\bar{c})$ is satisfied, the ϵs can be taken zero.

efficient firms in the subcoalition and the change for the $(k-l)$ inefficient firms. Now, under the zero transfer rule, inefficient firms always get the same transfer (namely zero), so that the effect of the manipulation on their pay-offs was, in fact, neutral. This leads us to consider, in the case of positive transfers, transfer structures that have the same characteristic as the zero transfer rule, viz. the transfers to be received by an inefficient firm are equal for all (relevant) states of nature. This is what we call the *partial anonymity property with respect to an inefficient firm.* With partially anonymous transfers, it does not matter for the inefficient firms in the subcoalition whether or not one of them is instructed to lie. For the efficient firms, the consequences of such a manipulation are, however, again negative: as soon as one inefficient firm lies, there will be more 'efficient' firms in the cartel, so that the cartel output (i.e. the monopoly output) must be divided among more firms. So, again, there is a negative externality on the efficient firms in the subcoalition. One is left to compare the direct gain for the inefficient firm with the negative externality exerted on the efficient firms. A sufficient condition for the overall benefit of manipulation to be negative is that transfers are taken such that the direct effect of misrepresentation is zero. This is obtained when the transfer that each (reportedly) efficient firm has to pay equals the benefit $\frac{l}{n-l}\{\pi^m(\underline{c}) - \Delta c q^m(\underline{c})\}$ that an inefficient firm gets when it is lying. We summarize this result in the following proposition:

Proposition 3 *A sufficient condition for the cartel to be implementable when collusion is possible is that the (partially anonymous) transfers are taken to be* $t(j) = \frac{n-j}{n}\{\pi^m(\underline{c}) - \Delta c q^m(\underline{c})\}$.

The rule of the above proposition is, in fact, collusion-proof with respect to all conceivable coalitions and dominant strategy incentive compatible. That a dominant strategy incentive compatible transfer rule is collusion-proof is a feature that this model has in common with other models of collusion, e.g. Laffont and Martimort (1997). Dominant strategy incentive compatibility implies that, without any compensation, it is costly for each firm not to tell the truth. This makes effective collusion more difficult to implement by the third party, in our model even impossible[18]. Still, it is easy to construct dominant strategy incentive compatibile rules that are not collusion-proof.

A dominant strategy incentive compatible rule naturally involves the use of more transfers than a Bayesian incentive compatible rule. It remains, therefore, to see whether the rule of the above proposition satisfies the participation constraint of the efficient type, the relevant upper-bound in this respect (by definition it satisfies the Bayesian incentive compatibility constraint of the efficient type). On the basis of two sufficient conditions and some simulations, we conclude that the (partially anonymous) dominant strategy incentive compatible transfer scheme

[18]Note that with this particular scheme of transfers neither $BIC(\underline{c})$ nor $BIC(\overline{c})$ is binding at the optimum. This implies that none of the incentive constraints in the program of the third party are binding at the optimum, so that the particular ϵs can be taken zero.

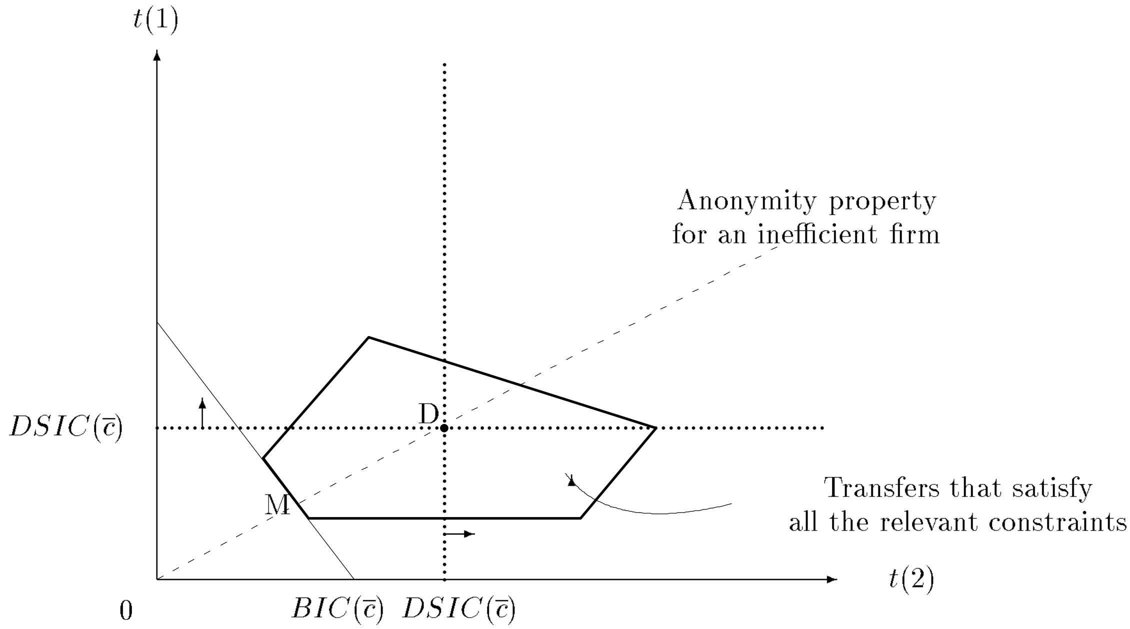

Figure 3: The two couples of transfers relating to the existence result.

$t(j) = \frac{n-j}{n}\{\pi^m(\underline{c}) - \Delta c q^m(\underline{c})\}$ implements the efficient cartel.

The rule highlighted above is sufficient to counter collusion, but not necessary; recall that it completely neutralizes the direct gain of the lying firms, so that the overall effect is strictly negative. There certainly exist transfer schemes that leave some gain to the lying firms but still produce an overall negative effect. Simulations[19]indicate that the set of transfers such that the anonymity property for an inefficient firm is satisfied and such that an inefficient firm is individually indifferent between lying and truth-telling to the cartel manager ($BIC(\overline{c})$ is binding) always enable the cartel manager to implement the efficient cartel (whenever positive transfers are necessary). A corollary is that the threat of collusion by subgroups seems to not to change the level of minimally required expected transfers used by the cartel manager to implement an efficient cartel.

In figure 3, two partially anonymous rules (the minimum expected transfer rule, M, and the dominant strategy rule, D) that can implement the efficient cartel in spite of the threat of subcoalition formation are illustrated for the case of a three firm industry.

[19]Available upon request.

7 Concluding remarks

In this paper we have addressed the question whether the approach followed in the existing literature to interpret the problem of cartel formation as a classical Principal-Agent problem is not too restrictive. In particular, we have verified whether one basic assumption underlying the use of the Revelation Principle, the assumption about subgroups of firms not attempting to jointly manipulate the cartel in their favour, is determinative for the outcome that a cartel can achieve. The principal result of the paper is that this is not the case in the context of a two type industry. Whenever a cartel can be implemented with transfers satisfying only the individual incentive and participation requirements of the cartel members, it is also possible to implement this cartel in a collusion-proof way, albeit with a more restricted set of transfers.

Several explanations can be given for this result. First of all, it must be noted that the degree of freedom in the choice of transfers is substantial in this two-type model (cf. d'Aspremont and Gérard-Varet (1979) and Mookherjee and Reichelstein (1992)). This freedom will surely help to find transfers that not only ensure individual implementability, but that also fight collusion. Furthermore, there is a strong congruence of interest between the cartel manager and the individual cartel members: after all, the cartel manager is acting costlessly on behalf of the members, by maximizing their total expected profits. A third observation is also intuitive: an inefficient firm that is instructed by the third party to represent itself to the cartel manager as being efficient exerts a negative externality on all firms that are truly efficient, including the efficient firms within the subcoalition. If, furthermore, the transfers to the inefficient firms do not depend on the number of inefficient firms present in the cartel and the direct gain of lying for the inefficient firms is not too large, then the total effect of the contemplated action is a negative one. In particular, as soon as the individual dominant strategy incentive compatibility constraint of an inefficient firm is satisfied, so will the group incentive compatibility constraints with respect to downward cost announcements.

In our model, collusion appears to have no bite. It is not yet clear to what extent the assumption of independence in the distribution of cost types influences this result. The literature leaves a somewhat mixed impression on this point. In Laffont and Martimort (1998), with two agents and correlated types, collusion by the grand coalition does matter.

As we have noted, the Collusion-Proofness Principle does not hold when the identity of the possibly colluding subcoalitions are not knowm to the cartel manager; it no longer needs to be the case that a cartel that is collusion-proof belongs to the class of optimal cartels. Nonetheless, the cartel contract that we have found to be optimal, is collusion-proof also with respect to simultaneously active subcoalitions. This result may be driven by the symmetry of the model[20].

Whether there are cartel contracts that implement the efficient cartel, while letting collusion

[20]There is both ex ante symmetry with respect to the cost levels and symmetry with respect to the objective

occur by (some) coalitions, remains an open question[21]. The literature has exhibited many settings in which it is in optimal for a principal to let collusion occur. In our context, the basic reason for the Collusion-Proofnesss Principle not to apply can be traced to the fact that, first, the cartel manager does not know the identity of the players he is confronted with (he knows the firms, but not the coalitions) and, secondly, he is not able to elicit this information.

A final remark can be made on the emphasis on the study of efficient cartels. What has been done in this model is to take the production vector of the cartel to be the efficient one (the first best solution) and then to see whether there exists a vector of transfers such that all BIR and BIC constraints are satisfied. As far as general transfers are concerned this is not a problem, as it is shown that an efficient cartel is indeed implementable. When transfers cannot be used ('weak cartels' in the meaning of McAfee and McMillan (1992)) it should be interesting to consider cartels that are not necessarily efficient. After all, a standard result of the mechanism design literature is that by distorting 'at the bottom' incentive compatibility constraints can be relaxed. What we could do, therefore, is to write the cartel manager's program in full, giving him also the opportunity to distort the production levels, and to see the impact of collusion in this case. This is for future research.

function of the cartel manager.

[21] At this point one may ask the question whether some subcoalitions are more likely to form than others, a question addressed in the literature on endogenous coalition formation (cf. Bloch (1995)). If the cartel manager could 'predict' which coalitions are most likely to form given the cartel scheme it is offering (e.g. because they are the only 'stable' coalitions), it might, in principle, be profitable to focus on these coalitions.

8 Appendices

8.1 Cournot competition under incomplete information

With Cournot competition under incomplete information a given firm i with cost c^i chooses its production quantity q^i given the expected total production of the other firms, $\mathbf{E}\{Q^{N-\{i\}}\}$: $\max_{q^i}(p(Q)-c^i)q^i = \max_{q^i}(1-\mathbf{E}\{Q^{N-\{i\}}\}-q^i-c^i)q^i$. The first-order condition for profit maximization is $q^i = \frac{1}{2}(1-\mathbf{E}\{Q^{N-\{i\}}\}-c^i) =: R^i(\mathbf{E}\{Q^{N-\{i\}}\})$, assuming that $c^i \leq 1-\mathbf{E}\{Q^{N-\{i\}}\}$, otherwise firm i will prefer to produce nothing. In a symmetric equilibrium, $\mathbf{E}\{Q^{N-\{i\}}\} = (n-1)\mathbf{E}\{q^i\} = (n-1)\frac{1}{2}(1-\mathbf{E}\{Q^{N-\{i\}}\}-\mathbf{E}\{c^i\})$, so that $\mathbf{E}\{Q^{N-\{i\}}\} = \frac{n-1}{n+1}(1-\mathbf{E}\{c^i\})$. Hence, the equilibrium quantities are $q^i_{\text{Cournot}}(c^i) = R^i(\mathbf{E}\{Q^{N-\{i\}}\}) = \frac{1}{2}(1-c^i-\frac{n-1}{n+1}(1-\mathbf{E}\{c^i\}))$, with expected profits $\mathbf{E}[\pi^i_{\text{Cournot}}(c^i)] = R^i(\mathbf{E}\{Q^{N-\{i\}}\}) = \frac{1}{4}(1-c^i-\frac{n-1}{n+1}(1-\mathbf{E}\{c^i\}))^2$. This results in the following profits:

$$\begin{cases} \mathbf{E}[\pi^i_{\text{Cournot}}(\underline{c})] & = \frac{1}{4(n+1)^2}(2(1-\underline{c})+(n-1)(1-\mu)\Delta c)^2 \\ \mathbf{E}[\pi^i_{\text{Cournot}}(\overline{c})] & = \frac{1}{4(n+1)^2}(2(1-\overline{c})-(n-1)\mu\Delta c)^2. \end{cases}$$

The above expressions hold under the requirement that $c^i \leq 1-\mathbf{E}\{Q^{N-\{i\}}\}, \forall i$, which translates into $\Delta c \leq \frac{2(1-\underline{c}}{2+(n-1)\mu} =: \Delta c_{\text{limit}}$. For $\Delta c \geq \Delta c_{\text{limit}}$, it is easily verified that only low cost firms will be active in equilibrium (will produce positive quantities) and that their expected profits are equal to $\frac{(1-\underline{c})^2}{(2+(n-1)\mu)^2}$.

8.2 Expressing the incentive and participation constraints in terms of bounds on expected total transfers

For a given transfer scheme t, the level of expected total transfers is given by $\mathbf{E}\{t(j)\} = \sum_{j=0}^{n}\binom{n}{j}\mu^j(1-\mu)^{n-j}t(j)$. One can rewrite this level as

$$\begin{aligned} \mathbf{E}\{t(j)\} &= n(1-\mu)\cdot\sum_{j=1}^{n-1}\binom{n-1}{j}\mu^j(1-\mu)^{n-1-j}\frac{1}{n-j}t(j) \\ &= \mathbf{E}\{t(j)\} = n\mu\cdot\sum_{j=0}^{n-1}\binom{n-1}{j}\mu^j(1-\mu)^{n-1-j}\frac{1}{j+1}t(j+1). \end{aligned}$$

8.2.1 The Bayesian incentive constraints

If the cartel manager wants to obtain truthful revelation of the private information by the firms he must satisfy

$$BIC(c^i): \mathbf{E}_{c^{N-\{i\}}}\{\pi^i(c^i,(c^i,c^{N-\{i\}}))\} \geq \mathbf{E}_{c^{N-\{i\}}}\{\pi^i(c^i,(\tilde{c}^i,c^{N-\{i\}}))\} \quad \forall(c^i,\tilde{c}^i)\in\{\underline{c},\overline{c}\}^2.$$

In the case of a symmetric and efficient cartel, the Bayesian incentive compatibility constraint for an efficient firm is

$$BIC(\underline{c}) : \sum_{j=0}^{n-1} \tbinom{n-1}{j}\mu^j(1-\mu)^{n-1-j}\frac{1}{j+1}(\pi^m(\underline{c}) - t(j+1)) \geq$$
$$\frac{1}{n}(1-\mu)^{n-1}(\pi^m(\overline{c}) + \Delta c q^m(\overline{c})) + \sum_{j=1}^{n-1} \tbinom{n-1}{j}\mu^j(1-\mu)^{n-1-j}\frac{1}{n-j}t(j).$$

Observing that $\sum_{j=0}^{n-1} \binom{n-1}{j}\mu^j(1-\mu)^{n-1-j}\frac{1}{j+1}t(j+1) = \sum_{j=1}^{n} \binom{n-1}{j-1}\mu^{j-1}(1-\mu)^{n-j}\frac{1}{j}t(j)$ we obtain by putting all the transfer terms $t(j)$ to one side

$$\sum_{j=0}^{n-1} \tbinom{n-1}{j}\mu^j(1-\mu)^{n-1-j}\frac{1}{j+1}\pi^m(\underline{c}) + (1-\mu)^{n-1}\frac{1}{n}(\pi^m(\overline{c}) + \Delta c q^m(\overline{c}))$$
$$\geq \sum_{j=1}^{n-1}((\tbinom{n-1}{j}\frac{1}{n-j}\mu + \tbinom{n-1}{j-1}\frac{1}{j}(1-\mu))\mu^j(1-\mu)^{n-1-j}t(j).$$

The right hand side equals $\sum_{j=1}^{n-1} \binom{n-1}{j-1}\mu^{j-1}(1-\mu)^{n-1-j}\frac{1}{j}t(j) = \sum_{j=0}^{n-1} \binom{n-1}{j}\mu^j(1-\mu)^{n-j}\frac{1}{j+1}t(j+1)$. It follows that an upperbound on the expected total transfer is given by

$$\mathbf{E}\{t(j)\} = n\mu \cdot \sum_{j=0}^{n-1} \tbinom{n-1}{j}\mu^j(1-\mu)^{n-1-j}\frac{1}{j+1}t(j+1)$$
$$\leq n\mu(1-\mu)[\frac{1}{n\mu}(1-(1-\mu)^n)\pi^m(\underline{c}) + (1-\mu)^{n-1}\frac{1}{n}(\pi^m(\overline{c}) + \Delta c q^m(\overline{c}))]$$

as $\sum_{j=0}^{n-1} \binom{n-1}{j}\mu^j(1-\mu)^{n-1-j}\frac{1}{j+1} = \sum_{j=1}^{n} \binom{n-1}{j-1}\mu^{j-1}(1-\mu)^{n-j}\frac{1}{j} = \sum_{j=1}^{n} \binom{n}{j}\mu^{j-1}(1-\mu)^{n-j}\frac{1}{n} = \frac{1}{n\mu}(1-(1-\mu)^n)$.

In a symmetric and efficient cartel, the Bayesian incentive compatibility constraint for an inefficient firm is

$$BIC(\overline{c}) : (1-\mu)^{n-1}\frac{1}{n}\pi^m(\overline{c}) + \sum_{j=1}^{n-1} \tbinom{n-1}{j}\mu^j(1-\mu)^{n-1-j}\frac{1}{n-j}t(j)) \geq$$
$$\sum_{j=1}^{n-1} \tbinom{n-1}{j}\mu^j(1-\mu)^{n-1-j}\frac{1}{j+1}(\pi^m(\underline{c}) - \Delta c q^m(\underline{c})) + t(j+1))$$

which, by a similar exercise as above, can be expressed as

$$\mathbf{E}\{t(j)\} \geq n\mu(1-\mu)[\tfrac{1}{n\mu}(1-(1-\mu)^n)(\pi^m(\underline{c}) - \Delta c q^m(\underline{c})) - (1-\mu)^{n-1}\tfrac{1}{n}\pi^m(\overline{c})].$$

8.2.2 The participation constraints

The participation constraints are as follows

$$BIR(c^i) : \mathbf{E}_{c^{N-\{i\}}}[\pi^i(c^i, (c^i, c^{N-\{i\}}))] \geq \mathbf{E}_{c^{N-\{i\}}}\{\pi^i_{\text{Cournot}}(c^i)\}, \forall i \in N.$$

In a symmetric and efficient cartel, the participation constraint for an efficient firm is

$$BIR(\underline{c}) : \sum_{j=0}^{n-1} \binom{n-1}{j} \mu^j (1-\mu)^{n-1-j} \frac{1}{j+1}[\pi^m(\underline{c}) - t(j+1)] \geq \mathbf{E}\{\pi^i_{\text{Cournot}}(\underline{c})\}$$

which translates into

$$\mathbf{E}\{t(j)\} \leq n\mu \cdot [\frac{1}{n\mu}(1-(1-\mu)^n)\pi^m(\underline{c}) - \mathbf{E}\{\pi^i_{\text{Cournot}}(\underline{c})\}].$$

Similarly, for an inefficient firm

$$BIR(\overline{c}) : \frac{1}{n}(1-\mu)^{n-1}[\pi^m(\overline{c})] + \sum_{j=1}^{n-1} \binom{n-1}{j} \mu^j (1-\mu)^{n-1-j} \frac{1}{n-j} t(j) \geq \mathbf{E}\{\pi_{\text{Cournot}}(\overline{c})\}$$

which becomes

$$\mathbf{E}\{t(j)\} \geq n(1-\mu) \cdot \{\mathbf{E}\{\pi_{\text{ Cournot}}(\overline{c})\} - \tfrac{1}{n}(1-\mu)^{n-1}\pi^m(\overline{c})\}.$$

8.3 Individual implementability of an efficient cartel

In order to show that it is always possible to find a transfer schemes t that satisfy all four individual constraints, we will compare the upper and lowerbounds on the total expected transfer $\mathbf{E}\{t(j)\}$ that are implied by these constraints.

8.3.1 Comparing $BIC(\underline{c})$ and $BIC(\overline{c})$

The upperbound on $\mathbf{E}\{t(j)\}$ implied by $BIC(\underline{c})$ is above the lowerbound on $\mathbf{E}\{t(j)\}$ implied by $BIC(\underline{c})$ if and only if

$$\frac{1}{n\mu}(1-(1-\mu)^n)\pi^m(\underline{c}) + (1-\mu)^{n-1}\frac{1}{n}(\pi^m(\overline{c}) + \Delta c q^m(\overline{c})) \geq$$
$$\frac{1}{n\mu}(1-(1-\mu)^n)(\pi^m(\underline{c}) - \Delta c q^m(\underline{c})) - (1-\mu)^{n-1}\frac{1}{n}\pi^m(\overline{c})$$

or $(1-(1-\mu)^n)q^m(\underline{c}) \geq \mu(1-\mu)^{n-1}q^m(\overline{c})$ which is the case as $1-(1-\mu)^n - \mu(1-\mu)^{n-1} = 1-(1-\mu)^{n-1} \geq 0$ and $q^m(\underline{c}) \geq q^m(\overline{c})$.

8.3.2 Comparing $BIR(\underline{c})$ and $BIR(\overline{c})$

In order to show that there exist transfers that simultaneously satisfy $BIR(\underline{c})$ and $BIR(\overline{c})$ it will proove to be useful to write the expected Cournot profits $\mathbf{E}\{\pi^i_{\text{Cournot}}(\underline{c})\}$ and $\mathbf{E}\{\pi^i_{\text{Cournot}}(\overline{c})\}$ using the summation operator: this will facilitate the making of term-by-term comparisons. Write $\mathbf{E}\{\pi^i_{\text{Cournot}}(\underline{c})\} = \sum_{j=0}^{n-1}\binom{n-1}{j}\mu^j(1-\mu)^{n-1-j}\cdot\underline{\pi}^i_{CN}(j+1)$, where $\underline{\pi}^i_{CN}(j+1)$ is the Cournot-Nash profit for a low cost firm when among the $(n-1)$ other firms in the industry, j firms turn out to be low cost. Similarly, express $\mathbf{E}\{\pi^i_{\text{Cournot}}(\overline{c})\}$ as $\sum_{j=0}^{n-1}\binom{n-1}{j}\mu^j(1-\mu)^{n-1-j}\overline{\pi}^i_{CN}(j)$, with $\overline{\pi}^i_{CN}(j)$ the Cournot profit for a high cost firm when there are j low cost firms among the remaining firms. Then the participation constraints for an efficient and inefficient firm can be written as

$$BIR(\underline{c}) : \sum_{j=0}^{n-1}\binom{n-1}{j}\mu^j(1-\mu)^{n-1-j}[\frac{1}{j+1}\pi^m(\underline{c}) - \frac{1}{j+1}t(j+1) - \underline{\pi}^i_{CN}(j+1)] \geq 0$$

$$BIR(\overline{c}) : \frac{1}{n}(1-\mu)^{n-1}\pi^m(\overline{c}) + \sum_{j=0}^{n-1}\binom{n-1}{j}\mu^j(1-\mu)^{n-1-j}[\frac{1}{n-j}t(j) - \overline{\pi}^i_{CN}(j)] \geq 0$$

Using $\frac{1}{j+1}\binom{n-1}{j} = \frac{1}{n}\binom{n}{j+1}$and $\frac{1}{n-j}\binom{n-1}{j} = \frac{1}{n}\binom{n}{j}$ and rearranging, these conditions become respectively:

$$\sum_{j=1}^{n-1}\frac{1}{n}\binom{n}{j}\mu^{j-1}(1-\mu)^{n-j}t(j) \leq \sum_{j=1}^{n}\binom{n-1}{j-1}\mu^{j-1}(1-\mu)^{n-j}[\frac{1}{j}\pi^m(\underline{c}) - \underline{\pi}^i_{CN}(j)]$$
$$\sum_{j=1}^{n-1}\frac{1}{n}\binom{n}{j}\mu^{j-1}(1-\mu)^{n-j}t(j) \geq \frac{1-\mu}{\mu}\sum_{j=0}^{n-1}\binom{n-1}{j}\mu^j(1-\mu)^{n-1-j}[\overline{\pi}^i_{CN}(j)] - \frac{1}{n}(1-\mu)^{n-1}\pi^m(\overline{c})$$

Hence, transfers can be found that satisfy the two constraints if

$$\sum_{j=1}^{n}\binom{n-1}{j-1}\mu^{j-1}(1-\mu)^{n-j}[\frac{1}{j}\pi^m(\underline{c}) - \underline{\pi}^i_{CN}(j)] \geq$$
$$\sum_{j=0}^{n-1}\binom{n-1}{j}\mu^{j-1}(1-\mu)^{n-j}[\overline{\pi}^i_{CN}(j)] - \frac{1}{n\mu}(1-\mu)^n\pi^m(\overline{c})$$

which becomes

$$\sum_{j=1}^{n-1}\binom{n-1}{j-1}\mu^{j-1}(1-\mu)^{n-j}\frac{1}{j}[\pi^m(\underline{c}) - j\underline{\pi}^i_{CN}(j) - (n-j)\overline{\pi}^i_{CN}(j)] +$$
$$\mu^{n-1}\frac{1}{n}[\pi^m(\underline{c}) - n\underline{\pi}^i_{CN}(n)] \geq \frac{1}{\mu}(1-\mu)^n[\overline{\pi}^i_{CN}(0) - \frac{1}{n}\pi^m(\overline{c})].$$

The above condition is met as $\pi^m(\underline{c}) - j\underline{\pi}^i_{CN}(j) - (n-j)\overline{\pi}^i_{CN}(j) \geq 0, \forall j = 0, ..., n$ and $\overline{\pi}^i_{CN}(0) - \frac{1}{n}\pi^m(\overline{c}) \leq 0$.

8.3.3 Comparing $BIC(\underline{c})$ and $BIR(\overline{c})$

After some adaptations, we can write $BIC(\underline{c})$ and $BIR(\overline{c})$ respectively as

$$\sum_{j=1}^{n-1}\frac{1}{n}\tbinom{n}{j}\mu^{j-1-j}(1-\mu)^{n-j}t(j)=$$

$$\sum_{j=1}^{n}\tbinom{n-1}{j-1}\mu^{j-1}(1-\mu)^{n-j}\frac{1}{j}\pi^m(\underline{c})+(1-\mu)^{n-1}\frac{1}{n}(\pi^m(\overline{c})+\Delta cq^m(\overline{c}))$$

and

$$\sum_{j=1}^{n-1}\frac{1}{n}\tbinom{n}{j}\mu^{j-1-j}(1-\mu)^{n-j}t(j)\geq\sum_{j=0}^{n-1}\tbinom{n-1}{j}\mu^{j-1}(1-\mu)^{n-1-j}[\overline{\pi}^i_{CN}(j)]-\frac{1}{n\mu}(1-\mu)^{n-1}\pi^m(\overline{c}).$$

We see that the upperbound implied by $BIC(\underline{c})$ is greater than the lowerbound implied by $BIR(\overline{c})$ if

$$\sum_{j=1}^{n}\tbinom{n-1}{j-1}\mu^{j-1}(1-\mu)^{n-j}\frac{1}{j}\pi^m(\underline{c})+(1-\mu)^{n-1}\frac{1}{n}(\pi^m(\overline{c})+\Delta cq^m(\overline{c}))\geq$$

$$\sum_{j=0}^{n-1}\tbinom{n-1}{j}\mu^{j-1}(1-\mu)^{n-1-j}[\overline{\pi}^i_{CN}(j)]-\frac{1}{n\mu}(1-\mu)^{n-1}\pi^m(\overline{c}).$$

Observe that with respect to the left hand side it holds that

$$\sum_{j=1}^{n}\tbinom{n-1}{j-1}\mu^{j-1}(1-\mu)^{n-j}\frac{1}{j}\pi^m(\underline{c})+(1-\mu)^{n-1}\frac{1}{n}(\pi^m(\overline{c})+\Delta cq^m(\overline{c}))=$$

$$\sum_{j=1}^{n}\tbinom{n}{j}\mu^{j-1}(1-\mu)^{n-j}\frac{1}{n}\pi^m(\underline{c})-(1-\mu)^{n-1}\frac{1}{n}\pi^m(\overline{c})+[(1-\mu)^{n-1}\frac{1}{n}\pi^m(\overline{c})-(1-\mu)^{n-1}\frac{1}{n}(\pi^m(\overline{c})+\Delta cq^m(\overline{c}))].$$

This amount is larger than $\sum_{j=1}^{n}\binom{n}{j}\mu^{j-1}(1-\mu)^{n-j}\frac{1}{n}\pi^m(\underline{c})-(1-\mu)^{n-1}\frac{1}{n}\pi^m(\overline{c})=(\frac{1-(1-\mu)^n}{\mu}-(1-\mu)^{n-1})\frac{1}{n}\pi^m(\underline{c})=\frac{1-(1-\mu)^{n-1}}{\mu}\frac{1}{n}\pi^m(\underline{c})$, where we have used that $\pi^m(\overline{c})\geq(\pi^m(\overline{c})+\Delta cq^m(\overline{c}))$. For the left hand side, it holds that

$$\sum_{j=0}^{n-1}\tbinom{n-1}{j}\mu^{j-1}(1-\mu)^{n-1-j}[\overline{\pi}^i_{CN}(j)]-\frac{1}{n\mu}(1-\mu)^{n-1}\pi^m(\overline{c})$$

$$=\sum_{j=1}^{n-1}\tbinom{n-1}{j}\mu^{j-1}(1-\mu)^{n-1-j}[\overline{\pi}^i_{CN}(j)]+\frac{1}{\mu}(1-\mu)^{n-1}\overline{\pi}^i_{CN}(0)-\frac{1}{n\mu}(1-\mu)^{n-1}\pi^m(\overline{c}).$$

This amount is smaller than $\sum_{j=1}^{n-1}\binom{n-1}{j}\mu^{j-1}(1-\mu)^{n-1-j}[\overline{\pi}^i_{CN}(j)]\leq\sum_{j=1}^{n-1}\binom{n-1}{j}\mu^{j-1}(1-\mu)^{n-1-j}[\overline{\pi}^i_{CN}(0)]=\frac{1-(1-\mu)^{n-1}}{\mu}\overline{\pi}^i_{CN}(0)$, where we have used that $\overline{\pi}^i_{CN}(0)\leq\pi^m(\overline{c})$ and that $\overline{\pi}^i_{CN}(j)\leq\overline{\pi}^i_{CN}(0),\forall j=1,...,n-1$. Now, as $\frac{1}{n}\pi^m(\underline{c})\geq\overline{\pi}^i_{CN}(0)$, we can conclude that the upperbound implied by $BIC(\underline{c})$ is greater than the lowerbound implied by $BIR(\overline{c})$.

8.3.4 Comparing $BIR(\underline{c})$ and $BIC(\overline{c})$ (partially involving simulations)

It is possible to find transfers for which the expected total level is above the lowerbound imposed by $BIC(\overline{c})$ and below the upperbound imposed by $BIR(\underline{c})$ if

$$n\mu(1-\mu)[\frac{1}{n\mu}(1-(1-\mu)^n)(\pi^m(\underline{c})-\Delta c q^m(\underline{c}))-(1-\mu)^{n-1}\frac{1}{n}\pi^m(\overline{c})] \leq$$
$$n\mu\cdot[\frac{1}{n\mu}(1-(1-\mu)^n)\pi^m(\underline{c})-\mathbf{E}\{\pi^i_{\text{Cournot}}(\underline{c})\}].$$

This inequality reduces to

$$\mathbf{E}\{\pi^i_{\text{Cournot}}(\underline{c})\} \leq \frac{1}{n}\pi^m(\underline{c})+\frac{1}{n\mu}(1-\mu-(1-\mu)^n)\Delta c q^m(\underline{c})$$
$$+(1-\mu)^n\frac{1}{n}(\pi^m(\overline{c})-\pi^m(\underline{c})+\Delta c q^m(\underline{c})). \quad (1)$$

Note first that it would be wrong to think that always $\mathbf{E}\{\pi^i_{\text{Cournot}}(\underline{c})\} \leq \frac{1}{n}\pi^m(\underline{c})$; when a low cost firm has a good probability to be the only efficient firm and the cost difference is fairly large, it may be that its conditionally expected Bayesian-Cournot profit is higher than the monopoly profit divided by all the firms.

Sufficient condition 1 *A sufficient condition for the lowerbound imposed by $BIC(\overline{c})$ to be below the upperbound imposed by $BIR(\underline{c})$ is that $\Delta c \leq \Delta c_{ES}$, where Δc_{ES} is the largest cost difference for which efficient firms are still willing to share the cartel profits equally with the inefficient firms.*

Observe first that the last two terms in inequality (1) are both positive (by a revealed preference argument, $\pi^m(\overline{c}) \geq \pi^m(\underline{c})-\Delta c q^m(\underline{c})$. A sufficient condition for inequality (1) to be satisfied is that $\mathbf{E}\{\pi^i_{\text{Cournot}}(\underline{c})\} \leq \frac{1}{n}\pi^m(\underline{c})$, i.e. low cost firms prefer sharing equally the cartel profit $\pi^m(\underline{c})$ to getting the expected profit out of Cournot competition under incomplete information. Elaborating on this, we obtain

$$\frac{1}{n}\pi^m(\underline{c}) \leq \mathbf{E}\{\pi^i_{\text{Cournot}}(\underline{c})\} \Leftrightarrow \frac{1}{n}\pi^m(\underline{c}) \leq \frac{1}{4(n+1)^2}(2(1-\underline{c})+(n-1)(1-\mu)\Delta c)^2$$
$$\Leftrightarrow \Delta c \leq \frac{(\sqrt{n}-1)^2(1-\underline{c})}{(n-1)\sqrt{n}(1-\mu)} =: \Delta c_{\text{ES}}.$$

The above calculation is based on the Cournot profit for an efficient firm when high costs firms are active in equilibrium (i.e. $\Delta c \leq \Delta c_{\text{limit}}$). Obviously, when $\frac{1}{n}\pi^m(\underline{c}) \leq \mathbf{E}\{\pi^i_{\text{Cournot}}(\underline{c})\}$ for $\Delta c \leq \Delta c_{\text{limit}}$, it is also the case for larger Δc. Therefore, when the obtained level Δc_{ES} turns out to be larger than Δc_{limit}, we can conclude that $\frac{1}{n}\pi^m(\underline{c}) \leq \mathbf{E}\{\pi^i_{\text{Cournot}}(\underline{c})\}$ for all Δc. As a corollary, an alternative sufficient condition that $n \geq \left(\frac{2-\mu}{\mu}\right)^2$. Indeed, when $\frac{1}{n}\pi^m(\underline{c}) \leq \mathbf{E}\{\pi^i_{\text{Cournot}}(\underline{c})\}$ for $\Delta c \leq \Delta c_{\text{limit}}$, it is also the case for larger Δc. Then, $\Delta c_{\text{ES}} \leq \Delta c_{\text{limit}} \Leftrightarrow \underline{c}+\frac{(\sqrt{n}-1)^2(1-\underline{c})}{(n-1)\sqrt{n}(1-\mu)} \leq \frac{2+(n-1)\mu\underline{c}}{2+(n-1)\mu} \Leftrightarrow n \geq \left(\frac{2-\mu}{\mu}\right)^2$.

Sufficient condition 2 *A sufficient condition for the lowerbound imposed by $BIC(\overline{c})$ to be below the upperbound imposed by $BIR(\underline{c})$ is that $\Delta c \geq \Delta c^0_{BIC}$, where Δc^0_{BIC} is the level of Δc for which, under zero transfers, a high cost firm is indifferent between lying and truthtelling.*

The zero transfer rule satisfies $BIC(\overline{c})$ if

$$0 \geq \frac{1}{n\mu}(1-(1-\mu)^n)(\pi^m(\underline{c}) - \Delta c q^m(\underline{c})) - (1-\mu)^{n-1}\frac{1}{n}\pi^m(\overline{c}).$$

Using the relations $\pi^m(.) = (q^m(.))^2$ and $q^m(\overline{c}) = q^m(\underline{c}) - \frac{1}{2}\Delta c$ this comes down to $\mu(1-\mu)^{n-1}\frac{1}{4}(\Delta c)^2 + (1-(1-\mu)^{n-1})\Delta c q^m(\underline{c}) - (1-(1-\mu)^{n-1})(q^m(\underline{c}))^2 \geq 0$. This relation with equality is quadratic in Δc and has two roots, one negative and one positive. From this, it follows that the inequality is met for $\Delta c \geq \Delta c^0_{BIC}$, where

$$\Delta c^0_{BIC} = \frac{1}{\mu(1-\mu)^{n-1}}\{\sqrt{(1-(1-\mu)^n)(1-(1-\mu)^{n-1})} - (1-(1-\mu)^{n-1})\}(1-\underline{c}).$$

Simulations Rewriting inequality (1), we get

$$\alpha(\Delta c)^2 + \beta q^m(\underline{c})\Delta c + \gamma(q^m(\underline{c}))^2 \leq 0, \tag{2}$$

where $\alpha = \frac{n(n-1)^2(1-\mu)^2}{4(n+1)^2} - \frac{1}{4}(1-\mu)^n$, $\beta = \frac{2n(n-1)(1-\mu)}{(n+1)^2} - \frac{1}{\mu}(1-(1-\mu)^{n-1})$ and $\gamma = \frac{4n}{(n+1)^2} - 1$. The determinant of relation (2) with equality is $(\beta^2 - 4\alpha\gamma)(q^m(\underline{c}))^2 > 0$ as $\gamma < 0$ for $n \geq 2$ and $\alpha \geq 0$. Furthermore, as $(\beta^2 - 4\alpha\gamma)(q^m(\underline{c}))^2 \geq \beta^2(q^m(\underline{c}))^2$, there is one positive and one negative root. Hence, for the bound of $BIC(\overline{c})$ to be below the bound of $BIR(\underline{c})$ one must have that $\overline{c} \leq \overline{c}'$, where

$$\overline{c}' = \underline{c} + \frac{1}{4\alpha}(-\beta + \sqrt{(\beta^2 - 4\alpha\gamma)})(1-\underline{c})$$

We know that if the bound of $BIC(\overline{c})$ is below the bound of $BIR(\underline{c})$ for $\overline{c} \leq \overline{c}_{\text{limit}}$, it also holds for all $\overline{c} \in [\overline{c}_{\text{limit}}, 1]$. Let us compare $\overline{c}'$ and $\overline{c}_{\text{limit}}$; if $\overline{c}' \geq \overline{c}_{\text{limit}}$, we can conclude that the bound of $BIC(\overline{c})$ is below that of $BIR(\underline{c})$ for all parameter values. Observe that both $\overline{c}'$ and $\overline{c}_{\text{limit}}$ are linear in $\underline{c}$ and that when $\underline{c} = 1$, we have $\overline{c}' = \overline{c}_{\text{limit}} = 1$. We therefore only have to compare $\overline{c}'$ and $\overline{c}_{\text{limit}}$ for $\underline{c} = 0$. Simple simulations on the resulting inequality in two parameters, n and μ, show that indeed $\overline{c}' \geq \overline{c}_{\text{limit}}$; we conclude that for all parameter values, one can find transfers that satisfy $BIC(\overline{c})$ and $BIR(\underline{c})$.

8.4 Minimal transfers

The two lowerbounds on the transfers are given by the constraints $BIC(\overline{c})$ and $BIR(\overline{c})$. We have

$$BIC(\overline{c}) \Leftrightarrow \frac{1}{n}\sum_{j=1}^{n-1}\tbinom{n}{j}\mu^{j-1}(1-\mu)^{n-j-1}t(j) \geq$$

$$-\frac{1}{n}(1-\mu)^{n-1}\pi^m(\overline{c}) + \frac{1-(1-\mu)^n}{\mu n}[\pi^m(\underline{c}) - \Delta c q^m(\underline{c})]$$

$$BIR(\overline{c}) \Leftrightarrow \frac{\mu}{n}\sum_{j=1}^{n-1}\tbinom{n}{j}\mu^{j-1}(1-\mu)^{n-j-1}t(j) \geq -\frac{1}{n}(1-\mu)^{n-1}\pi^m(\overline{c}) + \mathbf{E}\{\pi_{\text{Cournot}}(\overline{c})\}.$$

We will show that, whenever positive transfers are necessary to implement the efficient cartel, it is constraint $BIC(\overline{c})$ that determines the level of necessary expected transfers. This is equivalent to saying that

$$\frac{1-(1-\mu)^n}{n}[\pi^m(\underline{c}) - \Delta c q^m(\underline{c})] \geq \mathbf{E}\{\pi_{\text{Cournot}}(\overline{c})\} - \frac{1}{n}(1-\mu)^n\pi^m(\overline{c}) \tag{3}$$

for all parameter values for which $BIR(\overline{c})$ requires positive transfers. The latter requirement amounts to $(-(1-\mu)^{\frac{n-1}{2}} + \frac{2\sqrt{n}}{n+1} + \frac{(n-1)\sqrt{n}}{n+1}\mu)\Delta c \leq 2(-(1-\mu)^{\frac{n-1}{2}} + \frac{2\sqrt{n}}{n+1})q^m(\underline{c})$, i.e. to the requirement $\Delta c \leq \Delta c^0_{BIR}$, where Δc^0_{BIR} is the cost difference for which the zero transfer rule just satisfies $BIR(\overline{c})$

$$\Delta c^0_{BIR} = \frac{2(-(1-\mu)^{\frac{n-1}{2}} + \frac{2\sqrt{n}}{n+1})}{-(1-\mu)^{\frac{n-1}{2}} + \frac{2\sqrt{n}}{n+1} + \frac{(n-1)\sqrt{n}}{n+1}\mu}$$

where we can bear in mind that when $-(1-\mu)^{\frac{n-1}{2}} + \frac{2\sqrt{n}}{n+1} + \frac{(n-1)\sqrt{n}}{n+1}\mu < 0$, i.e. when μ is small enough, the zero transfer rule always satisfies $BIR(\overline{c})$, regardless of the cost difference.

Rewriting inequality (3) gives us $\alpha(\Delta c)^2 + \beta q^m(\underline{c})\Delta c + \gamma(q^m(\underline{c}))^2 \leq 0$, where $\alpha = n(2+(n-1)\mu)^2 - (n+1)^2(1-\mu)^n$, $\beta = 4(n-1)(n-1-2n\mu)$ and $\gamma = -4(n-1)^2$. If α is positive, i.e. $n(2+(n-1)\mu)^2 - (n+1)^2(1-\mu)^n > 0$ (μ not too small), then the discriminant Δ is positive and there is one positive and one negative root (which can be discarded). Call the relevant positive root $\Delta\hat{c}_+$. It holds that $BIC(\overline{c})$ requires higher transfers than $BIR(\overline{c})$ if for all $\Delta c \leq \Delta c^0_{BIR}$, we have that $\Delta c \leq \Delta\hat{c}_+$. By comparing $\Delta\hat{c}_+$ and Δc^0_{BIR} for the relevant parameter values, we can see whether this is the case. Observe that both Δc^0_{BIR} and $\Delta\hat{c}_+$ are linear in $\underline{c}$ and that when $\underline{c} = 1$, we have $\Delta c^0_{BIR} = \Delta\hat{c}_+ = 0$. We therefore only have to compare Δc^0_{BIR} and $\Delta\hat{c}_+$ for $\underline{c} = 0$. Simple simulations on the resulting inequality in two parameters, n and μ, show that indeed $\Delta c^0_{BIR} \leq \Delta\hat{c}_+$. When α is negative (μ not too large), it easily follows that $BIR(\overline{c})$ is satisfied by the zero transfer rule, so that also in this case $BIC(\overline{c})$ is necessarily the most demanding. We conclude that for all relevant parameter values, $BIC(\overline{c})$ requires higher transfers than $BIR(\underline{c})$.

8.5 The Collusion-Proofness Principle

Let us assume that the k firms of the subcoalition are the first k firms of the cartel. Let us consider a perfect Bayesian equilibrium of the overall game of a cartel contract CC offer in the presence of subcoalition formation such that a side mechanism $SM \neq SM_0$ is chosen by the third party.

CC maps the messages $m = (m^1, \cdots, m^k, \tilde{c}^{k+1}, \cdots, \tilde{c}^n) \in M^1 \times ... \times M^k \times \{\underline{c}, \overline{c}\}^{n-k}$ sent by the firms into an allocation $(q, t) \in \Re^n_+ \times \Re^n$. CC maximizes the cartel manager's welfare taking into account the continuation equilibrium of the game of coalition formation. We can restrict the space of messages for the members of the subcoalition to be $\{\underline{c}, \overline{c}\}^k$. SM is a side mechanism which can be taken as being a direct side mechanism mapping $\{\underline{c}, \overline{c}\}^k$ into the set of measures on the messages spaces. SM maximizes the sum of the firms' expected gains subject to Bayesian incentive constraints, budget balance conditions and Bayesian individual rationality constraints $\Pi^i(c^i) \geq \tilde{\Pi}^i(c^i)$ where $\Pi^i(c^i)$ is the expected gain of firm i when CC and SM are played.

Consider now the new cartel contract $CC' = CC \circ SM$. We shall prove that there exists a perfect Bayesian equilibrium of the overall game of cartel contract offer with coalition formation in which the cartel manager offers CC' which is a direct cartel contract from $\{\underline{c}, \overline{c}\}^n$ into the decision space $\Re^n_+ \times \Re^n$, the third party offers the null side-mechanism SM_0 and this choice is sustained with passive beliefs.

Because SM solves the third party's program with reservation gains $\tilde{\Pi}^i(c^i)$ the null side-mechanism SM_0 solves the third party's program with reservation gains $\Pi^i(c^i)$. Indeed, suppose that it is not the case. Then there would exist a side mechanism SM' such that the third party can achieve a strictly greater payoff for the subcoalition than with SM. Since by definition $\Pi^i(c^i) \geq \tilde{\Pi}^i(c^i)$ the third party's payoff from offering $SM' \circ SM$ would be strictly greater than that achieved with SM. But this would contradict that S is optimal when CC is offered. Hence, offering the cartel contract CC' ensures the cartel manager that there is a perfect Bayesian equilibrium of the overall game sustained with passive beliefs in which SM_0 is optimal from the point of view of the third party.

8.6 The third party's program

Because the Revelation Principle applies at the third party's level, in order to ensure the revelation of information and the participation of the firms in the subcoalition, the third party must solve the following problem

$$\max_{\{\phi,y\}} \mathbf{E}_{c^{N\setminus S}}\mathbf{E}_{c^S}\{\sum_{i\in S}\pi^i(c^i,\phi(c^i,c^{S-\{i\}}),c^{N\setminus S})\}$$

subject to

$$\begin{aligned}
BIC^{TP}(c^i)\ &:\ \mathbf{E}_{c^{N\setminus S}}\mathbf{E}_{c^{S-\{i\}}}\pi^i(c^i,\phi(c^i,c^{S-\{i\}}),c^{N\setminus S})\geq \\
&\qquad \mathbf{E}_{c^{N\setminus S}}\mathbf{E}_{c^{S-\{i\}}}\pi^i(c^i,\phi(\tilde{c}^i,c^{S-\{i\}}),c^{N\setminus S}),\ \forall(c^i,\tilde{c}^i)\in\{\underline{c},\overline{c}\}^2\,\forall i\in N \\
BIR^{TP}(c^i)\ &:\ \mathbf{E}_{c^{N\setminus S}}\mathbf{E}_{c^{S-\{i\}}}\pi^i(c^i,\phi(c^i,c^{S-\{i\}}),c^{N\setminus S})\geq\tilde{\Pi}^i(c^i),\ \forall c^i\in\{\underline{c},\overline{c}\}\,\forall i\in N \\
BB^{TP}(c^S)\ &:\ \sum_{i\in S}y^i(c^S)=0,\ \forall c^S\in\{\underline{c},\overline{c}\}^k
\end{aligned}$$

where $\tilde{\Pi}^i(c^i)$ is the gain of firm i when it rejects the side-mechanism proposed by the third party and plays non-cooperatively the symmetric cartel contract proposed by the cartel manager with passive beliefs. Now, let us denote by $\underline{\delta}_i^{TP}$, $\overline{\delta}_i^{TP}$, $\underline{\nu}_i^{TP}$, $\overline{\nu}_i^{TP}$ and $\rho^{TP}(c^S)$ the Lagrange multipliers associated respectively to $BIC_i^{TP}(\underline{c})$, $BIC_i^{TP}(\overline{c})$, $BIR_i^{TP}(\underline{c})$, $BIR_i^{TP}(\overline{c})$ and $BB^{TP}(c^S)$. Optimizing with respect to the side-transfers $y^i(c^S)$ and $y^j(c^s)$, where firm i is efficient and firm j is inefficient we find the following conditions

$$\begin{aligned}
\rho(c^S)&=\mu^l(1-\mu)^{k-l}[\underline{\delta}_i^{TP}+\underline{\nu}_i^{TP}-\overline{\delta}_i^{TP}] \\
\rho(c^S)&=\mu^l(1-\mu)^{k-l}[\overline{\delta}_j^{TP}+\overline{\nu}_j^{TP}-\underline{\delta}_j^{TP}].
\end{aligned}$$

Immediate computations give us then the optimality conditions for the manipulation function.

8.6.1 The collusion-proofness constraints for an efficient cartel

Writing the conditions such that identity function is the optimal manipulation function for the third party gives us all the collusion-proofness constraints.

$$\begin{aligned}
CPC^k(0,l')\,\forall 0<l' : (1-\mu)^{n-k}\{&\frac{k}{n}[\pi^m(\overline{c})-\overline{\epsilon}\frac{\mu}{1-\mu}\Delta cq^m(\overline{c})]- \\
&[\pi^m(\underline{c})-\Delta cq^m(\underline{c})-\overline{\epsilon}\frac{\mu}{1-\mu}\Delta cq^m(\underline{c})-t(l')+\frac{k-l'}{n-l'}t(l')]\}+ \\
&+\sum_{j=1}^{n-k}\alpha_{n-k}(j)\{\frac{k}{n-j}t(j)-\frac{l'}{l'+j}[\pi^m(\underline{c})-\Delta cq^m(\underline{c})- \\
&\qquad \overline{\epsilon}\frac{\mu}{1-\mu}\Delta cq^m(\underline{c})-t(l'+j)]-\frac{k-l'}{n-l'}t(l'+j)\}\geq 0
\end{aligned}$$

$$CPC^k(l,0): \sum_{j=0}^{n-k} \alpha_{n-k}(j)\{\frac{l}{l+j}[\pi^m(\underline{c}) + \underline{\epsilon}\frac{1-\mu}{\mu}\Delta c q^m(\underline{c}) - t(l+j)] + \frac{k-l}{n-l-j}t(l+j)\} -$$
$$\sum_{j=1}^{n-k} \alpha_{n-k}(j)[\frac{k}{n-j}t(j)] - (1-\mu)^{n-k}\{\frac{k-l}{n}[\pi^m(\overline{c}) - \overline{\epsilon}\frac{\mu}{1-\mu}\Delta c q^m(\overline{c})]$$
$$+ \frac{l}{n}[\pi^m(\overline{c}) + \Delta c q^m(\overline{c}) + \underline{\epsilon}\frac{1-\mu}{\mu}\Delta c q^m(\overline{c})]\} \geq 0$$

$$CPC^k(l,l')\,\forall l > l': \sum_{j=0}^{n-k} \alpha_{n-k}(j)\{\frac{l}{l+j}[\pi^m(\underline{c}) - t(l+j) + \underline{\epsilon}\frac{1-\mu}{\mu}\Delta c q^m(\underline{c})]$$
$$+ \frac{k-l}{n-l-j}t(l+j) - \frac{l'}{l'+j}[\pi^m(\underline{c}) + \underline{\epsilon}\frac{1-\mu}{\mu}\Delta c q^m(\underline{c}) - t(l'+j)] - \frac{k-l'}{n-l'-j}t(l'+j)\} \geq 0$$

$$CPC^k(l,l')\,\forall l < l': \sum_{j=0}^{n-k} \alpha_{n-k}(j)\{\frac{l}{l+j}[\pi^m(\underline{c}) - t(l+j) + \underline{\epsilon}\frac{1-\mu}{\mu}\Delta c q^m(\underline{c})] +$$
$$\frac{k-l}{n-l-j}t(l+j) - \frac{l}{l'+j}[\pi^m(\underline{c}) + \underline{\epsilon}\frac{1-\mu}{\mu}\Delta c q^m(\underline{c}) - t(l'+j)] -$$
$$\frac{l'-l}{l'+j}[\pi^m(\underline{c}) - \Delta c q^m(\underline{c}) - \overline{\epsilon}\frac{\mu}{1-\mu}\Delta c q^m(\underline{c}) - t(l'+j)] - \frac{k-l'}{n-l'-j}t(l'+j)\} \geq 0$$

where we denote for clarifying purposes $\alpha_{n-k}(j) = \binom{n-k}{j}\mu^j(1-\mu)^{n-k-j}$.

8.7 A sufficient condition for the zero transfer rule to be collusion-proof

Under the zero transfer rule, coalitions of firms never have interest to report to be less efficient than they are in reality: $CPC^k(l,0) \Leftrightarrow \sum_{j=0}^{n-k} \alpha_{n-k}(j)[\frac{l}{l+j}\pi^m(\underline{c})] \geq \frac{1}{n}(1-\mu)^{n-k}[k\pi^m(\overline{c}) + l\Delta c q^m(\overline{c})]$ where $\alpha_{n-k}(j) = \binom{n-k}{j}\mu^j(1-\mu)^{n-k-j}$. The left hand side of this inequality is a sum of positive terms. For $j=0$ the corresponding term is equal to $(1-\mu)^{n-k}\pi^m(\underline{c})$ which is greater than the right hand side of the inequality. Hence, $CPC^k(l,0)$ is satisfied by the zero transfer rule. For $l > l'$, $CPC^k(l,l') \Leftrightarrow \sum_{j=0}^{n-k} \alpha_{n-k}(j)[\frac{l}{l+j} - \frac{l'}{l'+j}]\pi^m(\underline{c}) \geq 0$, which is obviously satisfied.

Now, we will show that $CPC^k(0,k)$ implies $CPC^k(0,l')$, $\forall l' < k$. Indeed, $CPC^k(0,l')$ where $l' < k$:

$$\sum_{j=0}^{n-k} \alpha_{n-k}(j)\frac{l'}{l'+j}[\pi^m(\underline{c}) - \Delta c q^m(\underline{c})] \leq (1-\mu)^{n-k}\frac{k}{n}\pi^m(\overline{c}).$$

Obviously, as $\frac{l'}{l'+j}$ is increasing in l', the above constraint is most demanding for $l' = k$. Last, because $(1-\mu)^{n-k}\frac{k}{n}$ is increasing in k, the individual constraint $BIC(\overline{c})$ is the most demanding among the constraints $CPC^k(0,k)$.

It remains to prove that $CPC^k(0,l')$ implies all the constraints $CPC^k(l,l')$, where $l' > l$. The zero transfer rule satisfies $CPC^k(l,l')$ where $l' > l$ if

$$\sum_{j=0}^{n-k} \alpha_{n-k}(j)[\frac{l}{l+j}\pi^m(\underline{c}) - \frac{l'}{l'+j}\pi^m(\underline{c}) + \frac{l'-l}{l'+j}\Delta c q^m(\underline{c})] \geq 0$$

i.e.

$$\sum_{j=0}^{n-k} \alpha_{n-k}(j)\frac{l'}{l'+j}[\pi^m(\underline{c}) - \Delta c q^m(\underline{c})] \leq \sum_{j=0}^{n-k} \alpha_{n-k}(j)[\frac{l}{l+j}\pi^m(\underline{c}) - \frac{l}{l'+j}\Delta c q^m(\underline{c})].$$

Hence, $CPC^k(0,l')$ is more demanding than $CPC^k(l,l')$ if

$$(1-\mu)^{n-k}\frac{k}{n}\pi^m(\overline{c}) \leq \sum_{j=0}^{n-k} \alpha_{n-k}(j)\{\frac{l}{l+j}\pi^m(\underline{c}) - \frac{l}{l'+j}\Delta c q^m(\underline{c})].$$

As $\pi^m(\underline{c}) - \frac{l}{l'}\Delta c q^m(\underline{c}) \geq (1-\frac{l}{l'})\pi^m(\underline{c}) \geq \frac{1}{l+1}\pi^m(\underline{c})$, and $\pi^m(\overline{c}) = [q^m(\overline{c})]^2 = [q^m(\underline{c}) - \frac{1}{2}\Delta c]^2$ we obtain the following very sufficient condition:

$$(1-\frac{k}{n})q^m(\underline{c})^2 - (\frac{l}{l+1} - \frac{k}{n})\Delta c q^m(\underline{c}) - \frac{k}{4n}(\Delta c)^2 \geq 0.$$

A sufficient condition is then that this inequality is satisfied for $\Delta c = q^m(\underline{c})$, or

$$k(l+1) \leq 4n \, \forall l = 1, ..., k-1 \Leftrightarrow k \leq 2\sqrt{n}.$$

8.8 A sufficient condition for a cartel to be collusion-proof

Consider the transfer scheme $t(j) = \frac{n-j}{n}\pi^m(\underline{c}) - \Delta c q^m(\underline{c})$. It satisfies the collusion-proofness constraints $CPC^k(l,l')$ and $CPC^k(l',l)$, with $l > l'$ if

$$\sum_{j=0}^{n-k} \alpha_{n-k}(j)[(\frac{l'}{l'+j} - \frac{l}{l+j})\pi^m(\underline{c})] \leq$$
$$\sum_{j=0}^{n-k} \alpha_{n-k}(j)[(\frac{k-l}{n-l-j} - \frac{l}{l+j})t(l+j) + (\frac{l'}{l'+j} - \frac{k-l'}{n-l'-j})t(l'+j)] \leq$$
$$\sum_{j=0}^{n-k} \alpha_{n-k}(j)[(\frac{l'}{l'+j} - \frac{l}{l+j})\pi^m(\underline{c}) + \frac{1}{l+j}(l-l')\Delta c q^m(\underline{c})].$$

Using the fact that the transfers chosen here are such that $t(l+j) = (n-l-j)t(n-1)$, $nt(n-1) = \pi^m(\underline{c}) - \Delta c q^m(\underline{c})$ and that $l > l'$ it is immediate to check that these two inequalities are always satisfied. Indeed, the first inequality becomes equivalent to $\sum_{j=0}^{n-k} \alpha_{n-k}(j)\frac{-j(l-l')}{(l+j)(l'+j)}\pi^m(\underline{c}) \leq \sum_{j=0}^{n-k} \alpha_{n-k}(j)[\frac{k(l+j)-ln}{l+j} - \frac{k(l'+j)-l'n}{l'+j}]t(n-1) \Leftrightarrow \sum_{j=0}^{n-k} \alpha_{n-k}(j)\frac{j(l-l')}{(l'+j)(l+j)}\Delta c q^m(\underline{c}) \geq 0$, which is obviously satisfied. With the same kind of computations, we find that the second inequality is equivalent to $\sum_{j=0}^{n-k} \alpha_{n-k}(j) \frac{l'(l-l')}{(l+j)(l'+j)} \Delta c q^m(\underline{c}) \geq 0$ which is also satisfied under our starting assumptions. Note that the revealed preference argument applied respectively to an efficient and an inefficient firm when a symmetric and efficient cartel is implemented gives us $\pi^m(\underline{c}) \geq \pi^m(\bar{c}) + \Delta c q^m(\bar{c})$ and $\pi^m(\bar{c}) \geq \pi^m(\underline{c}) - \Delta c q^m(\underline{c})$.

Note that the proposed transfer scheme satisfies the requirement of dominant strategy implementation. Under the partial anonymity property, an inefficient firm receives the same transfer in each state of nature: $\frac{1}{n-j}t(j) = \frac{1}{n-j-1}t(j+1)$ or, equivalently, $t(j) = \frac{n-j}{n-1}t(1)$. Dominant strategy implementation amounts to

$$\frac{1}{j+1}[\pi^m(\underline{c}) - \Delta c q^m(\underline{c})] \leq \frac{1}{n-j}t(j) + \frac{1}{j+1}t(j+1) \leq \frac{1}{j+1}\pi^m(\underline{c}).$$

8.9 Implementability of an efficient cartel

We now check whether the collusion-proof transfer scheme $t(j) = \frac{n-j}{n}[\pi^m(\underline{c}) - \Delta c q^m(\underline{c})]$ satisfies the participation constraint of an efficient firm (the relevant upperbound in this respect). The scheme satisfies $BIR(\underline{c})$ if

$$\mathbf{E}\{\pi^i_{\text{Cournot}}(\underline{c})\} \leq \frac{1}{n}\pi^m(\underline{c}) + \frac{1-\mu}{n\mu}(1-(1-\mu)^{n-1})\Delta c q^m(\underline{c}). \tag{4}$$

Sufficient condition 1 *A sufficient condition for the collusion-proof transfer scheme $t(j) = \frac{n-j}{n}\pi^m(\underline{c}) - \Delta c q^m(\underline{c})$ to satisfy the participation constraint of an efficient firm is that $\Delta c \leq \Delta c_{ES}$, where Δc_{ES} is the largest difference for which efficient firms are still willing to share the cartel profits equally with the inefficient firms.*

A sufficient condition for inequality (4) to hold is that $\mathbf{E}\{\pi^i_{\text{Cournot}}(\underline{c})\} \leq \frac{1}{n}\pi^m(\underline{c})$, which gives $\Delta c \leq \Delta c_{\text{ES}}$ (see Section 8.3.4).

Sufficient condition 2 *A sufficient condition for the collusion-proof transfer scheme $t(j) = \frac{n-j}{n}(\pi^m(\underline{c}) - \Delta c q^m(\underline{c}))$ to satisfy $BIR(\underline{c})$ is that the cost difference is large enough: $\Delta c \geq q^m(\underline{c})$*

Simulations Rewriting inequality (4), we get

$$\alpha(\Delta c)^2 + \beta q^m(\underline{c})\Delta c + \gamma(q^m(\underline{c}))^2 \leq 0 \tag{5}$$

where $\alpha = n(n-1)^2\mu\ (1-\mu)^2$, $\beta = 8n(n-1)\mu(1-\mu)^n$ and $\gamma = 16n\mu - 4\mu(n+1)^2$. As $\gamma < 0$, the determinant of relation (5) with equality is positive. Furthermore, as $(\beta^2 - 4\alpha\gamma)(q^m(\underline{c}))^2 \geq \beta^2(q^m(\underline{c}))^2$, there is one positive and one negative root. Call the positive root Δc^*. Then, for the collusion-proof transfer scheme $t(j) = \frac{n-j}{n}(\pi^m(\underline{c}) - \Delta c q^m(\underline{c})$ to satisfy $BIR(\underline{c})$ it is enough that for all relevant Δc (i.e. $\Delta c \leq \Delta c_{\text{limit}}$), it holds that $\Delta c \leq \Delta c^*$. This will be the case whenever $\Delta c^* \geq \Delta c_{\text{limit}}$. Observe that both Δc^* and Δc_{limit} are linear in $\underline{c}$ and that when $\underline{c} = 1$, we have $\Delta c^* = \Delta c_{\text{limit}} = 0$. We therefore only have to compare Δc^* and Δc_{limit} for $\underline{c} = 0$. Simple simulations on the resulting inequality in two parameters, n and μ, show that indeed $\Delta c^* \geq \Delta c_{\text{limit}}$.; we conclude that for all parameter values, the collusion-proof transfer scheme $t(j) = \frac{n-j}{n}(\pi^m(\underline{c}) - \Delta c q^m(\underline{c}))$ satisfies $BIR(\underline{c})$.

References

[1] Bloch, F. (1995) “Endogenous structures of association in oligopolies”, *Rand Journal of Economics*, 26: 537-56.

[2] Cramton, P. and T. Palfrey (1990), “Cartel enforcement with uncertainty about costs”, *International Economic Review*, 31: 17-47.

[3] Cramton, P. and T. Palfrey (1995), “Ratifiable mechanisms: learning from disagreement”, *Games and Economic Behaviour*, 10(2): 255-283.

[4] D’Aspremont, C. and L.A. Gérard-Varet (1979), “Incentives and incomplete information”, *Journal of Public Economics*, 11: 25-45.

[5] Eckbo, P. (1976), “The future of world oil”, Ballinger, Cambridge, Massachussets USA.

[6] Kihlstrom, R. and X. Vives (1989), “Collusion by asymmetrically informed duopolists”, *European Journal of Political Economy*, 5: 371-402.

[7] Kihlstrom, R. and X. Vives (1992), “Collusion by asymmetrically informed firms”, *Journal of Economics and Management Strategy*, 1: 371-396.

[8] Laffont, J.-J. and D. Martimort (1997), “Collusion under asymmetric information”, *Econometrica*, 65(4): 875-911.

[9] Laffont, J.-J. and D. Martimort (1998), “Mechanism design under collusion and correlation”, *mimeo.*

[10] McAfee, R.P. and J. McMillan (1986), “Bidding rings”, *American Economic Review*, 82: 579-599.

[11] Mookherjee, D. and S. Reichelstein (1992), “Dominant strategy implementation of Bayesian incentive compatible allocation rules”, *Journal of Economic Theory*, 56: 378-399.

[12] Phlips, L. (1995), “Competition policy: a game-theoretic perspective”, Cambridge University Press, Cambridge, United Kingdom.

[13] Roberts, K. (1985), “Cartel behaviour and adverse selection”, *Journal of Industrial Economics*, XXXIII(4): 401-413.

[14] Selten, R. (1973), “A simple model of imperfect competition where four are few and six are many”, *International Journal of Game Theory*, 2: 141-201.